

INGRID HOLMBERG

EFFECTS OF SOME
TRIALS TO IMPROVE
MATHEMATICS TEACHING

Lund 1975

CWK GLEERUP



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41552412_0073

Effects of some trials to improve mathematics teaching

AV

INGRID HOLMBERG

FIL. LIC., KLM

AKADEMISK AVHANDLING

som för avläggande av filosofie doktorsexamen vid
samhällsvetenskapliga fakulteten vid universitetet i Lund
kommer att offentligen försvaras å Lärarhögskolan i Malmö,
sal E 122, fredagen den 23 maj 1975, klockan 10

INGRID HOLMBERG

Effects of some trials to improve mathematics teaching



CWK GLEERUP



STUDIA PSYCHOLOGICA ET PAEDAGOGICA
SERIES ALTERA XXVI

INGRID HOLMBERG

Effects of some trials to improve mathematics teaching

CWK Gleerup is the imprint for the scientific and scholarly publications
of LiberLäromedel Lund.

Contents

Acknowledgements 7

Introduction 9

1. *Methods studies—investigations of Effects 1 and 2* 22
 - 1.1 *Background and aims* 22
 - 1.1.1 Examples of experiments with programmed teaching 23
 - 1.1.2 Examples of experiments to measure attitudes towards programmed teaching 25
 - 1.1.3 Examples of experiments to measure classroom behavior 26
 - 1.2 *Problems* 27
 - 1.3 *Methods* 27
 - 1.3.1 Experimental design 27
 - 1.3.2 Subjects of experiments 28
 - 1.3.3 The role of the teacher 30
 - 1.3.4 Teaching methods 30
 - 1.3.5 Measuring instruments 31
 - 1.3.5.1 Mathematics tests 31
 - 1.3.5.2 Reading test 32
 - 1.3.5.3 Observation schedule 32
 - 1.3.5.4 Form H: an instrument for the measurement of the subject preferences of schoolchildren 33
 - 1.3.5.5 Essay writing—Attitude Test H 34
 - 1.3.5.6 Question-frequency notation and instruction-comprehension test for measuring working habits 35
 - 1.3.6 Processing methods 36
 - 1.4 *Results and discussion* 36
 - 1.4.1 Analysis of changes in achievement in mathematics 36
 - 1.4.2 Analysis of changes in achievements in reading ability 37
 - 1.4.3 Analysis of behavior assessments 38
 - 1.4.4 Analysis of preference for mathematics (Form H) 40

1.4.5	The attitude towards machine teaching as reflected in essay-writing or measured by Attitude test H	41
1.4.6	Analysis of working habits	43
1.4.7	Some studies of relations	44
1.4.7.1	Mathematics achievement—preference for mathematics	44
1.4.7.2	Mathematics achievement—reading ability	44
1.4.7.3	Correct answers in the test papers—correct answers in the series in the programmed material	45
1.5	<i>Summary</i>	45
1.6	<i>Conclusions</i>	47
2.	<i>Comparison of curricula: investigation of Effects 3</i>	50
2.1	<i>Background and aims</i>	50
2.2	<i>Problems</i>	55
2.3	<i>Methods</i>	55
2.3.1	Experimental design and methods of analysis	55
2.3.2	Subjects	56
2.3.3	Teachers	57
2.3.4	Study materials and activities	58
2.3.5	Instruments	58
2.3.5.1	Number conception tests	58
2.3.5.2	Knowledge test in mathematics	59
2.3.5.3	Tests of attitudes towards mathematics: Attitude MA	60
2.4	<i>Results and discussion</i>	61
2.4.1	Analysis of number conception	61
2.4.2	Analysis of knowledge in mathematics	61
2.4.3	Analysis of attitudes towards mathematics	62
2.5	<i>Summary</i>	63
2.6	<i>Conclusions</i>	64
3.	<i>Final discussion</i>	65
	<i>References</i>	73

Acknowledgements

These three studies have been conducted at the Department of Educational and Psychological Research at the Malmö School of Education.

The head of the Department, Professor Åke Bjerstedt, was the scientific leader of the two first studies, those concerning programmed learning in mathematics. This part of the thesis would probably not have come into existence in its present form without his advice on matters both small and great and his careful examination of the manuscript in all its stages. I express my sincere gratitude to him.

In the third study, a curriculum comparison, concentrating on the content of mathematics courses, my scientific leader was Professor Olof Magne. His specialist knowledge of mathematics has been most valuable, and I appreciate it.

I wish to thank Docent Bernt Larsson for good advice on the statistical analyses. I am also much obliged to Docent Gunnar Bergendal for giving me his views on the development of mathematics. My sincere thanks to Fil. kand. Agneta Sternerup, who did the data computing operations, Gillian Nilsson, B.A. for translating this thesis into English, Mr Lennart Wramborg for drawing the figures and Mrs Karin Tommeschat who has been an excellent secretary of my projects.

The Educational authorities in Malmö and Klippan kindly allowed me to carry out the experiments in their schools. The cooperation of 60 teachers and their pupils in different schools made it possible to collect data. I would like in particular to thank Lektor Arne Strelert and his classes, who for six terms tried teaching machines in their mathematics lessons.

The support from my colleagues at the department of Mathematics, Malmö School of Education, has been most valuable. Fil. mag. Nils Sander and Mrs Anne-Marie Anfelt prepared the mathematics tests.

These studies have partly been financed by a grant from the Swedish Board of Education.

Introduction

During the past decennium pedagogues all over the world have taken a keen interest in school mathematics. Today's science requires knowledge of modern mathematics. A thorough review has been made of the goals, content and methods of school mathematics. This thesis presents three investigations of effects in mathematics carried out in Sweden. The first two deal with problems of methods and are two longterm experiments each lasting three terms, using the new methodological bases of the 1960's: Programmed material presented via teaching machines. Since so much emphasis was placed on individualization in the 1962 curriculum for the comprehensive school (Lgr 62), it was considered extremely desirable that new methods for individualized mathematics teaching should be tried out. The investigations were carried out in grades 7—9 of the comprehensive school and were focussed on the differences in achievement, attitudes and student behavior that can conceivably emerge between groups that have been taught by different methods. The third study of effect consists of a curriculum comparison and concentrates on the content of mathematics courses. It analyzes and compares the effects of teaching mathematics according to two different curricula, Lgr 62 and Lgr 69, with regard to the development of the concept of numbers, knowledge of mathematics and attitudes to the subject. The investigations have been presented in the following reports:

- Holmberg, I. Effekten av programmerad undervisning på vissa prestations-, attityd- och beteendevariabler, jämfört med konventionell undervisning. Matematik — högstadiet. /The effect of programmed teaching in some achievement, attitude and behavior variables, compared to conventional teaching. Mathematics—secondary level./ *Pedagogisk-psykologiska problem* (Malmö, Sweden: School of Education), No. 18, 1965. (Rev. ed., 1975.)
- Holmberg, I. Maskinundervisning kombinerad med lärarledd smågruppsundervisning, jämfört med konventionell undervisning. /Mechanical teaching combined with teacher-led teaching in small groups, compared to con-

ventional teaching./ *Pedagogisk-psykologiska problem* (Malmö, Sweden: School of Education), No. 30, 1965. (Rev. ed., 1975.)

Holmberg, I. Effekter av ny läroplan i matematik (Lgr 69) på vissa aspekter av talbegreppets utveckling, kunskaper i matematik och attityder mot ämnet. /Effects of a new curriculum in mathematics (Lgr 69) on certain aspects of the development of the concept of numbers, knowledge in mathematics and attitudes towards the subject./ *Pedagogisk-psykologiska problem* (Malmö, Sweden: School of Education), No. 236, 1974.

This introduction first presents the discussion of methods that took place when the method experiments were being carried out and then describes the development of mathematics and some of the large research projects on which the revision of the curriculum was based.

The *debate on methods* during the 1960's was centered on programmed teaching, teaching machines and computer-based learning. The great expansion in education together with the lack of teachers were the driving forces behind the demand for measures to rationalize the educational sector and make it more effective. Much of the research done during the 1960's has concerned the problem of the individualization of teaching, particularly in mathematics. The National Society for the Study of Education and the Association for Supervision and Curriculum Development USA, published year-books with the title "Individualizing Instruction" in both 1962 and 1964. In the attempts to produce material that could meet the needs of the individual student, programmed material was developed.

The programmed material, which can be described as detailed teaching plans, was intended to lead to a pre-determined goal. In this material certain laws of learning psychology are used, such as systematic reinforcement techniques. The awareness of having succeeded in an assignment, of having got something right, is sufficiently satisfying to serve as reinforcement. In order to create a good learning effect, it is necessary to create a systematic pattern of small reinforcements. In programmed instruction the aim is to arrange an optimum learning climate, in which each individual works actively, at his own rate, with a low percentage of errors and an immediate experience of succeeding. A program of information and tasks is built up in such easy steps that the number of mistakes is always kept low. The program can be presented by a teaching machine. Such a machine can be constructed in various ways, but it must fulfil certain requirements. It must present information and tasks in such a way that solutions and the rest of the program are concealed. The student can construct his own answer or be asked to choose between different alternative answers. After the student has given his answer, the machine presents the correct answer and the next step is started. The programmed material can serve as a complete form of instruction, replacing a teacher.

One of the foremost scientific advocates of programmed instruction and teaching machines during the 1950's and 1960's was B. F. Skinner, who claimed that the reinforcement conditions themselves in an organism's environment are the factor that most influences learning. Modern research into the process of learning had earlier had little effect on methods. But working from experimental behavior analyses of animal experiments, Skinner formulated the principles that have been used in programming teaching material:

1. divide up the information into small units and present one at a time
2. provide reinforcement after each correct response
3. raise the demands successively so that the organism is forced to render the entire desired behavior in order to obtain further reinforcement.

Skinner's programs are most often linear, which means that the student always advances in the program. In a Skinner program the student constructs his own answer and all the students usually take the same sequence through the program (Skinner, 1954). In many ways programmed teaching has been the starting point for the development of educational technology. Thus Skinner has indirectly influenced the origin of a modern method of planning and implementing education as a whole.

While Skinner preferred to present new information in small units, thereby diminishing the risk of getting incorrect answers, N. A. Crowder advocated another variant of program, large units of information followed by different alternative answers (Crowder, 1963). Depending on which answer one chooses, one gets varying instruction and is steered through the program, which is a so-called branching program. The main characteristic of the method is that weak students are given a more detailed and longer schooling by being given instruction concerning their incorrect answers, while the better students are permitted to advance more quickly. The programs are presented both in book form, so-called "scrambled books" and mechanically. Crowder's branching model presupposes more complicated machines than Skinner's linear programs do. The most advanced teaching machines make use of computers to store results and choose program sequences. Suppes (1966) describes three types of computer-based instruction (Computer Assisted Instruction, CAI), namely "drill and practice systems," "tutorial systems" and "dialogue systems." In "drill and practice systems" the student is given tasks of a training nature and the programs are sequenced with regard to error frequency and answering speed. The "tutorial system" is more complicated. The programs are stored in the computer on film and audio-tapes. The student is given tasks via ear-phones and microfilm pictures on a television screen. The student-constructed answers can be written on a keyboard or with a "light pen." In multiple-choice answers the student uses a system of pushbuttons. The computer chooses the

program according to the quality of the answer. The most advanced interaction student computers are tested in "dialogue systems," which exist as elementary prototypes and which assume that the student learns a computer language in order to be able to communicate with the computer. Perhaps the most elaborate dialogue at this time has been produced by Starkweather, who has developed an "author language": Starkweather's PILOT /Programed Instruction, Learning or Teaching/ (Starkweather, 1965). Man and machine are regarded as a cooperating unit—a cybernetic system.

Teaching machines were no new phenomenon in the '60's. As early as 1926 Sidney Pressey of Ohio State University had published a report on "A simple apparatus which gives tests and scores—and teaches." This machine had programs based on alternative answers. It gave the correct answers and had a counting mechanism that noted wrong answers (Pressey, 1926). Pressey and his disciples did important pioneering work, but their methods were no immediate pedagogical success. Teaching machines of this kind were used, however, in military training in the Allied Armed Forces as early as 1942 (Davies, I. J. & Hartley, J. 1972, p. 1). At the end of the '50's the Russian sputniks appeared and accelerated a newly-awakened interest in educational questions in general and teaching machines in particular in the USA. L. E. Homme and Robert Glaser commented in 1959 on the abundance of machines, "this ridiculous situation in which machines outnumber programs" (Davies, I. J. & Hartley, J. 1972, p. 1).

The interest in the machines themselves cooled during the '60's, while books and booklets became common forms of presenting programmed material. This material was admittedly time-consuming and expensive in production but was nevertheless valuable, since it contributed to solving what was the most acute school problem in many countries—the lack of teachers.

UNESCO backed programmed teaching in under-developed countries with projects in Jordan and Nigeria among other places.

The principles formulated by Skinner on the basis of his reinforcement theories have come to be applied in a number of different fields. Their application within education is naturally the most important and most comprehensive but in clinical psychology too, ideas and methods have emerged in recent years under Skinner's influence. This applies primarily in the field of "Behavior modification," which concerns the treatment of certain psychological problems. Skinner's program has been used successfully in training autistic children and in the treatment of catatonia.

At the beginning of the '60's the question was put: "Which gives the best learning results, conventional teaching or programmed?" At the end of the '60's the question had been modified: "Which combination of teachers and other staff with different study material is best for achieving the study goal of the students concerned?" The latter question has led to the development

of instructional method-material systems and to more detailed studies of their different qualities and effects—the development of study materials is systematized. This development—using principles for the production of programmed material in developing other products—can be noted in the surveys concerning programmed instruction by Leith, 1969 and Deterline et al., 1970. Bjerstedt describes this development: “The focus has changed from a more limited interest area (a particular type of product) to a broader field: programmers are now more interested in the process, and the products of the process are now conceptualized to encompass different media and media combinations” (Bjerstedt, 1972, p. 13). Some of the journals of programmed instruction extended their titles to include a broader term; for example, “Programmed Learning and Educational Technology,” London and “Programmiertes Lernen, Unterrichtstechnologie und Unterrichtsforschung,” Berlin.

From the programming of teaching was developed educational technology. It was presented in Sweden in 1966 and became an academic subject in 1971. “Educational technology is the branch of pedagogics that deals with the development and application of systematic methods for the planning, implementation and evaluation of teaching, making use of experiences from a number of sociological and behavioral disciplines” (Course folder I in educational technology, TRU/Hermods, 1971).

In 1966 educational technology was anchored in Skinner’s learning theories, with a behavior-centered view of human change and development. Goals were formulated for the teaching and expressed in operational terms, i.e. it was stated what facts the student should know or how he should behave in a problem situation. The goal formulations were placed in relation to the group of students for whom the teaching was to be designed—a student analysis was made. The student analysis implied taking individual differences into consideration and it was necessary to decide what content would make it possible to achieve the goals—analysis and sequencing of subject-matter was carried out. The next stage in the planning concerned the choice of media and methods. Attempts were made to arrange the learning situation as effectively as possible, with forms of presentation that aroused the interest of the students and feedback that raised the students’ motivation. A procedure of successive evaluations was necessary. The study material was tested continually while it was being constructed and on the basis of this testing the material was revised. In the terminal field testing, data were collected so that it would be possible to state program effects in a manual. The mathematics program that has been used in my first two investigations of effect has been produced in accordance with the technical model outlined here.

During the 1970’s there has been a shift in educational technology from the behaviorist viewpoint towards cognitive ideas, following Piaget (1969) and

Bruner (1970). Teaching goals are given a less definitive content—they are not described in such exact terms as before, but rather in structures expressing a development. Questions concerning evaluation are also treated in a different way than before, since the entire educational process is regarded as a development steered by the activities of the participants.

Cognitive models have also been used in designing computer-based information with so-called “adaptive machines.” Gordon Pask (1969) emphasizes that there must be cooperation between what he calls “the external teaching mechanism” and “the internal teaching mechanism”—between teaching and learning.

Programs have been developed with a control hierarchy of a cybernetic model. Representatives of both pedagogics and psychology and subject experts have cooperated in writing the programs. They have tried to use as many media as possible for the giving of information and the registration of results—the multi-media approach.

Today the video-recorder, computer and tape-recorder all have their accepted place in the use of programmed material. This is now used for various kinds of occupational training, especially for specific parts of the training in which exact goal behaviors can be decided, e.g. in medical training, office training, shipbuilding, radar training (both civil and military).

Within the school sector programmed material is available in most subjects, but is used infrequently—the greatest frequency of use is in mathematics and statistics. Advanced programs with computer-based information are used only in a few places in the world, mostly in the USA. In Sweden the computer has been introduced experimentally into the upper secondary school, the emphasis there being placed on letting the students try out the possibilities offered by the computer, particularly in technical subjects. The students have been trained to write and run simple programs in the programming language BASIC. Björk (1974) describes how the work with BASIC has influenced the numeric and algebraic skills of the students and changed their attitudes towards mathematics positively.

The many experiments with “teacher-replacing” appliances, programmed material and computer-based learning during the '60's have not produced a single perfect method of teaching mathematics, but they have reduced the blind faith in technology as such and indicated possibilities of employing technology in the complex process that is learning.

The development of mathematics: Mathematics is ancient both as a science and a school subject. Nevertheless its development during the twentieth century, both quantitatively and qualitatively, has surpassed all that has happened during earlier centuries. Descriptions of the world around us are given increasingly frequently as mathematical models. This has led to a growing need for comprehension of mathematical contexts.

Mathematical theories have often been complete long before their application. That has been the pattern throughout the development of mathematics. Differential and integral calculus first took shape in the 17th century. An important driving force then was the study of dynamics. Newton and Leibniz naturally did not foresee the use to which their theories would be put, e.g. within the theory of wireless telegraphy (cf. Hall, p. 135).

A remarkable amount of today's applied mathematics has its origin in the distant past. Basic algebra, the use of logarithms and the calculus of probability have their roots in the sixteenth and seventeenth centuries. Later they have been further developed and applied in different fields. Topology goes back to the eighteenth century (Euler). Cantor's formulations of the theory of sets at the end of the nineteenth century has had an effect on the further development of topology. In the eighteenth century too Lagrange took the first steps towards the theory of groups in his investigations of algebraic equations. Galois put forward in the nineteenth century a complete theory of groups. In modern physics the theory of groups is indispensable. In the nineteenth century Gauss studied among other things curved surfaces. His studies were continued by Riemann and Minkowski, who laid the geometrical foundations for Einstein's theory of relativity.

These few examples alone indicate that the mathematics of our time has been built up by the contributions of many mathematicians over a long period of time. The range of knowledge has widened during the twentieth century, but it is based on the work of our predecessors. The question is not so much who have founded a certain part of mathematics, but rather in which way have different mathematicians contributed to the development of their science?

One of the main features of modern mathematics is its abstraction and the high demands made on stringency. "Earlier mathematical problems usually arose from descriptions of nature. Now problems are formulated from a purely logical viewpoint. Curves became first functions, then ordered pairs of numbers and finally mappings. Symmetry rotations in geometry became abstract groups, dice games and card games turned into set functions in abstract outcome space. Triangle geometry has developed into a study of properties, which remain unchanged throughout continuous deformations-topology. In this development there has been a powerful integration of different areas." (Hall, 1970, p. 134.)

The development of mathematical statistics during the past hundred years through the work of Pearson and Fischer has also provided mathematical aids for the behavioral sciences. Today educational research is dependent not only on the calculation of probability but also on modern numeric aids. Increasingly extensive data material can be analyzed as machines take over the processing. The most recent events within mathematics are not reflected directly in school mathematics. On a longterm basis, however, the development of

mathematics has produced a new view of its possibilities as an aid in various contexts, such as new prerequisites for the content of school mathematics. The efforts made to bring the school's mathematics teaching into step with modern requirements have come to the fore almost simultaneously in many of the more developed countries and the developmental work concerning course content and methods that has preceded the new curricula has an international stamp. Begle (1970) uses the word "revolution" in describing international developments in school mathematics during the years 1950—70. In the public debate this development has in recent years been associated with the introduction of the theory of sets into the school courses. The intention of the theory of sets, however, has primarily been to utilize it as a common language for the different parts of school mathematics: arithmetic, algebra, geometry, statistics, theory of probability, theory of function etc. Different kinds of sets, such as number sets, function sets and point sets, in which concepts such as order, continuity, distance and measurement are studied, form the foundation of modern analysis.

For a long time algebra was regarded as some kind of generalized arithmetic. Nobody imagined that an algebraic system could exist with a structure other than the arithmetic one. Gradually, however, mathematical systems were developed that differed structurally from ordinary arithmetic. In the nineteenth century, for example, George Boole developed an algebra of sets. When calculating with numbers, addition and multiplication are the basic operations, while in Boolean algebra one uses set operations. Boolean algebra can be utilized in such widely separated areas as the theory of electrical circuits and analysis of laws in jurisprudence.

The goals of school mathematics now aim more and more at an understanding of structures and relations. This has led to changes in the content, e.g. the stronger accentuation of numeric laws, a geometry based on the translation axiom and the introduction of counting with bases other than the base of ten.

Many mathematic items are dealt with at an earlier age than before. Counting with positive numbers and negative numbers, the number line, counting with different number bases and powers, solution of equations, for example, are now brought in during grades 4—6. In grades 7—9 new operations such as vector calculation and trigonometry are added. The theory of function is emphasized and an extended course in statistics is given.

In the train of the new content of school mathematics followed changes in *methods and material*. A large number of research projects in different countries have tried to solve the problems of mathematics methodology and develop teaching material. They have restructured the teaching process, tried out new ideas on how mathematics should be taught, tested different measures for individualization and shown the importance of group work. The content of

the mathematics courses has been developed into a mathematics based on understanding. Some larger projects have attracted particular attention.

School Mathematics Study Group (SMSG), the largest mathematics project in the USA, conducted during the years 1962—1967 a comprehensive longitudinal study of students from grades 4—12, the so-called National Longitudinal Study of Mathematical Abilities (NLSMA). The aim was to develop mathematics courses and study material and to improve syllabuses. Variables such as attitudes, achievements and teacher data were analyzed by means of multivariate methods. A large amount of data was collected and the results have been presented in thirtyodd reports. Some results from the SMSG material are described in Begle (1972). SMSG students were often superior to students using conventional material in understanding mathematical concepts, but inferior in numerical counting. The NLSMA reports illustrate that the evaluation of curricula is a complicated task and that few clear generalizations can be made on the basis of the results.

The Madison project in the USA, led by Robert B. Davies, developed a new curriculum in mathematics for the 8—16 age-group, with a selection of course content that followed the ideas of the Piaget school. The material is not specified for particular grades. The project organizes learning environments in which students can study essential aspects of our culture and learn to function in it. Methodologically “experience lessons” and “discussion lessons” are suggested. During “experience lessons” the students work with different kinds of laboratory material, such as Cuisenaire rods, Geo-boards, Tic-tac-toe. Thereby they investigate mathematical structures themselves and discover mathematical relations. In the “discussion lessons” the teachers and students discuss the relations and theories that the students have discovered during the “experience lessons.”

British experiences, represented by e.g. the Nuffield project using the “discovery method,” have greatly influenced the teaching methods of the Madison project. Computer-based teaching with PLATO IV is also on the program. This model of computer-mediated teaching implies a dialogue between student and computer. It is thought that low-achievement students can make better use of their mental capacities by using concrete study material. Younger students have the greatest advantage of the material. They obtain better achievements than one would have expected considering their age. In all age-groups the students increase their ability to understand complex mathematical problems without losing counting skills. The project makes use of video-recordings of student behavior in assessing the system. There is no systematic comparison with other forms of teaching (Barson, Cochran & Davies, 1970). Today the Madison project has a well-developed training program for teachers, which is used above all in Illinois, New York and California.

The University of Illinois Committee on School Mathematics (UICSM) first

attracted attention by its research on mathematics teaching in high school. During more recent years "The Arithmetic Project" at the same university has produced new ideas on how to develop the teaching process, i.e. work with mathematics in the classroom in grades 1—6. The project does not test curricula but method-material systems, which are intended to increase creativity in mathematics lessons. The teachers are trained in what is called "Open education," i.e. utilizing all possible stimuli in the work, such as films, discussions and newspapers. They are encouraged to vary the work by means of individualization, group work and class teaching. No systematic evaluation of the project has been presented.

In England the Nuffield Foundation, London, has focussed on material for in-service training of teachers in mathematics (Nuffield Mathematics Teaching Project). In connection with this in-service training they have produced student material for the 5—13 age-group. Piaget's principles are taken into consideration and the methods that are stressed are independent work by the students, experiments and discussions. The material is being evaluated internally.

IMU (Individualized mathematics teaching), a Swedish research and development project 1963—1973, analyzed material, methods and organization—three factors that together form the IMU system. The goals of the project were presented by its initiator, Curt Öreberg, in the following way:

- to construct and test self-instructional study material in mathematics
- to test suitable teaching methods and work forms for the use of this material
- to try out different ways of grouping students and making use of the teachers in order to achieve maximal effect of the material and methods
- to measure by means of the material constructed the effects of individualized teaching (in connection with a comparison with conventional teaching).

In IMU's activities the emphasis was placed on grades 7—9 of the comprehensive school. Some experiments have been carried out in grades 4—6, and the upper secondary school. Since students in grades 7—9 worked in unstreamed classes, which led to the organizational discontinuation of division into alternative courses, it was possible to group the students more flexibly and make use of teacher teams.

The IMU students mostly work independently with newly constructed self-instructional material under teacher guidance. The IMU material was expensive to produce, but the cost of the teaching was kept down by letting the students work in large groups and with lower teacher density. In addition teacher assistants were employed, cheaper labor with office training, as a complement to the teacher. Hereby the project studied the effect of different organizational models.

The methodological intention was that the IMU student should alternate between independent work and group teaching—a methodological parallel to the methods of the machine class in investigation of effects 2 (Holmberg, 1966). The intentions of the IMU project concerning teacher tutoring and group teaching were not carried out, however, at all the experimental schools, partly due to the cheeseparing of the state authorities, which resulted in far too narrow economic limits. Rationalization was throughout the guiding principle of the IMU system.

The main results of the investigation of effects were presented in Larsson (1973 a and b): The students accepted the individual method of working but desired more group teaching. The individual method of working appears to favor the high-achievement students but to some extent to disfavor the low-achievers, who are considered to be more reliant on the teacher. The basic attitudes of the teachers towards individualized mathematics teaching are reasonably positive. They express dissatisfaction, however, with the students' formal treatment of the tasks, which is considered to be worse when they work with the individual method.

When the rate of work is completely individualized, certain organizational consequences follow. The data of the material study show that at the end of grade 8, about 60 % of the students had kept up a normal rate of study. During the latter part of IMU's experimental period, some steering of the students' rate of work occurred. A minimum rate was introduced for students who worked slowly, but the fast students were very seldom held back in their work.

Students who have chosen the more difficult course in mathematics know more—these students represent the top intellectual group. Because of the differences in the content of courses, it has not been possible to compare the knowledge of the IMU students with that of students who had worked with other study material.

The size of the class had little effect on the knowledge and attitudes of the students. The role of the teacher was changed with the IMU material. Contact with individual students increases, while contacts with the whole class decreases. Planning and administrative tasks increase.

The experiences gained from the IMU project, both the negative and the positive ones, aroused an intense debate on the forms for mathematics teaching. This debate was of significance for the construction of the new curriculum and the design of study materials in mathematics in Sweden. The IMU project is the largest Swedish mathematics venture that has received international publicity.

The Nuffield Resources for Learning Project had part of the IMU material translated and tested in England. Michael Williams points out the effectivity of IMU in the Times Educational Supplement 5.10.73: "The children's initial

reaction was one of great enthusiasm . . . We were particularly impressed by the attitude and progress of weaker pupils . . . The most able pupils worked well and naturally responded best when personally encouraged . . . When the year ended, we were faced with yet another problem, that of fitting a suitable second-year course to a highly unusual first-course. However, although our supplies of materials were exhausted, we were sufficiently impressed with the whole exercise to start adapting the IMU scheme for use with our new first-year intake."

The IMU material is being tried out in Norway since 1968. The students there enjoy working with the material on the whole. The teachers consider that the system encourages cooperation between the students—weak students are helped by more knowledgeable classmates. For the faster students the IMU system provides a better teaching situation than conventional teaching (Meltingsaether, 1972).

Yet another Swedish research project is working on a design for mathematics teaching for the upper level of the comprehensive school—the Lund system of mathematics. The aim here is for the students to be able to work independently. Together with the teacher the student should be able to choose material according to his results in the recurrent diagnostic tests. The material consists of textbooks, workbooks of varying degrees of difficulty, diagnostic tests, test papers, audiotapes and transparencies for the overhead projector. The scientific leader of the project, R. Flinck, reports a positive attitude on the part of the students towards the Lund system's methods of work (Flinck, 1970). The attitudes of the teachers towards the system is reported as being relatively positive, but they consider that the system gives them more work than the conventional teaching methods they have used previously. Varying class sizes have been tried out in the Lund system, and the system proved to function better in conventional classes with about 30 students than in large classes. Here the results diverge from those obtained in the IMU experiments. The IMU experiments show small differences between varying class sizes. The material of the Lund system has now been completely revised and is used in classes of normal size. The final reports on the project are anticipated around the middle of the '70's.

"The large projects have led to changes in the curricula of several countries, including Sweden" (Werdelin, 1973, p. 97). "In Sweden MUNK (Matematikundervisningens Nordiska Kommitté = the Nordic Committee on Mathematics Teaching), NKMM (Nordiska Kommittén för Modernisering av Matematikundervisningen = the Nordic Committee for the Modernization of Mathematics Teaching) and IMU have provided guidelines for the content of the curriculum," states the MUT group in mathematics in their memorandum of 22.10.74 (MUT = Målskrivning och utvärdering = Goals and evaluation). The NKMM, appointed in 1960, had conducted extensive experiments with experi-

mental texts during the '60's. These introduced new elements such as the theory of sets, vectors and the theory of probability, while at the same time removing certain geometrical items. The experimental texts were tried out in both the comprehensive school and the upper secondary school with good results (Nordiska utredningar, 1967: 9). MUNK, appointed in 1968, worked on the basis of NKMM's report to design in more detail methods for mathematics teaching in grades 1—6. The new Swedish curriculum was ready in 1969 and in the autumn of 1970 grades 1, 4 and 7 of the comprehensive school started working with the new curriculum. The syllabus was introduced gradually and by 1973 the entire comprehensive school was working with it. Not until 1979 will all students have worked with the '69 curriculum throughout their entire school career. When the curriculum was introduced, work started immediately on the Following-up of the Curriculum in the Comprehensive School (LUG = LäroplansUppföljning i Grundskolan). The third part of this thesis is a contribution to this follow-up.

1. Method studies—investigations of Effects 1 and 2

This summary is based on two reports:

1. Holmberg, I. Effekten av programmerad undervisning på vissa prestations-, attityd- och beteendevariabler, jämfört med konventionell undervisning. (Matematik — Högstadiet.) /Programmed and conventional classroom instruction, compared with respect to certain achievement, attitude and behavior variables./ *Pedagogisk-psykologiska problem* (Malmö, Sweden: School of Education), No. 18, 1965. (Rev. ed., 1975.)
2. Holmberg, I. Maskinundervisning kombinerad med lärarledd smågruppsundervisning, jämfört med konventionell undervisning. /Programmed instruction combined with teacher-supervised small-group instruction, compared to conventional classroom instruction./ *Pedagogisk-psykologiska problem* (Malmö, Sweden: School of Education), No. 30, 1966. (Rev. ed., 1975.)

1.1 Background and aims

This project, which was conducted at our institute during the years 1962—1965, was initiated by the great interest felt by pedagogues in the opportunities opened up for the comprehensive school by the use of programmed material. Behind the experiments using programmed study lay a desire to organize teaching in such a way as to make the optimum use of the teaching body. Teaching machines and programs should free the teacher for more qualified tasks, such as individual tutoring.

The programming of the teaching material created prerequisites for an automatation of the teaching and a consequent saving in teacher resources. In our first investigation the intention was not to rationalize the teacher out of existence but to see how much the students could learn on their own from a program and see the effects of the program on the students and the learning situation. The teacher was instructed to adopt a passive role in the machine class. Our intention was to evaluate and compare the learning situation and

the learning of mathematics continuously in two groups of students: one with conventional teacher-supervised teaching and the other with programmed teaching and teaching machines, what is known as automated instruction. The investigation focussed on the differences in achievements, attitudes and student behavior that can result from these two forms of teaching in secondary school classes. During three terms, 1962—1963, the effect of a maximum use of teaching machines was compared with the effect of conventional teaching.

The results of these first experiments inspired a new investigation of effects in which the methodology was modified. The students were now offered an alternation of individual work with programmed material and teacher-led instruction in small groups. The teacher was given the opportunity of developing contacts and cooperation with the students in small-group teaching, thus gaining a more central and active place in the learning process. Our aim in this new experiment was that programmed teaching should complement the teacher rather than replace her/him.

In the first experiment a difference in the working habits was noted between students working with programmed material and other students. Therefore a closer study was made in the second investigation of effects of the working habits of different classes. This new investigation was focussed on differences in achievements, attitudes and student behavior in secondary school classes which can result from working with programmed material and teaching machines (of the Skinner type) combined with teacher-led small-group teaching on the one hand and working without teaching machines on the other.

1.1.1 *Examples of experiments with programmed teaching*

Few methods for the presentation of information material have been the subject of so much research during the '60's as methods involving programming. Skinner, who anchored programming in learning psychology, required of an effective program:

- that the student should work at his own speed
- that the material should be presented in small steps
- that the material should be well-sequenced
- that the student should actively process the material
- that the student should receive reinforcement—receive the correct answer immediately. (Skinner, 1954.)

These requirements have been tested scientifically.

Sequence studies had been made, mostly with short programs. Strangely enough in some of the experiments, no significant difference in learning could be shown between students who used programs with the frames in a random

sequence (Zuckerman, Marshall & Groesberg, 1961; Roe, K. V., 1962; Levin & Baker, 1963). In longer programs there was found to be a greater need for an ordered sequence and the importance of this varied between different types of material (Feldhusen, 1963).

Small steps—small percentage of errors: “Smaller steps mean among other things better focussing of what is essential, more opportunities for activation and greater facility in arranging suitable sequences” (Bjerstedt, 1963, p. 166). Evans, Glaser & Homme (1960) gave four alternative versions of a program with steps of differing length. The program with most frames and the smallest steps produced the best learning results. Coulson & Silberman (1960) obtained similar results in their investigations.

In a series of experiments with film material, it was found that the learning result improved if the length of the steps in the program was gradually increased than if one worked with constantly short or long steps (Maccoby & Sheffield, 1958; Margolius & Sheffield, 1961; Weiss, Maccoby & Sheffield, 1961).

The learning effect of programs with *student-constructed answers* has been compared to the effect of programs with readymade multi-choice answers. Most investigations show no significant difference in learning between programs with student-constructed answers and programs with multi-choice answers (Gropper & Lumsdaine, 1961; Stolurow & Walker, 1962 et al.).

Immediate confirmation of the correctness of the answer (feedback) has proved to be essential for learning—most investigations have found significantly more learning when there is immediate feedback of the right answer. The possibility has also been studied of giving “feedback” only after a series of frames, e.g. after 5 or 10 frames, which form a unit. Glaser & Taber (1961) found no significant difference in learning with 100 % reinforcement of answers and 50 or 20 % reinforcement. They point out that the need for reinforcement is greater when the percentage of errors is high. When the percentage of mistakes is kept low, as in a typical linear program, immediate reinforcement is less essential.

Individual rate of work does not produce the most effective learning for all students. Maccoby & Sheffield (1958) proved in a study that only the best students learned more, when they were allowed to work at their own rate. This is probably a consequence of the nature of the material.

In addition to these principles of Skinner, Crowder’s technique with *branching programs* has also been tested scientifically. Roe (1962) compares 7 different types of branching. He finds no significant differences either in learning time or in the final level of achievement between students reading all the frames in the linear program and those that skip some sections. Nor does any extra form of revision lead to superior achievements.

The learning effect of programmed material compared to other teaching

methods is still relatively uncertain, despite the numerous investigations that have been reported. Most of these studies show no significant difference in learning by different methods, according to a research survey by Silberman (1962). In most cases only short experiments have been involved—often no more than a few hours' work with programmed material (Hughes & McNamara, 1961). Here it has been possible to show the so-called Hawthorne effect—that which is new is found stimulating. It is impossible to determine the extent to which this effect has influenced the results. Hough (1962) compared machine teaching with conventional teaching in a relatively extensive experiment, a three-week course for students of education. He found that machine teaching was equally effective as conventional class teaching as far as the amount of material learned was concerned. If the control group was not given the chance of reading their notes during their spare time, there was a significant difference in effectivity in favor of machine teaching. In this experiment machine teaching saved time compared to conventional teaching. The effect that the initial enthusiasm can have had is not known, however. The only way of escaping the Hawthorne effect is probably to test the programmed material during a longer period—one year or more.

In year-long experiments with learning machines and programs in English (spelling, vocabulary etc.) most comparisons report no significant difference in learning between machine classes and control classes with conventional class teaching (Porter, 1959).

One long experiment conducted in Sweden, the Jönköping experiment, trying out material in mathematics for grade 7, produced no definite answers (Lane, 1963). Strict comparison of the results for the machine classes and control classes with conventional teaching could not be made, namely, since the programmed material could not be produced at the same rate as it was consumed.

1.1.2 *Examples of experiments to measure attitudes towards programmed teaching*

Attitudes towards programmed teaching are often reported as being very positive, especially in American studies (Skinner & Holland, 1960; Eigen, 1963). Short experiments with programs are a popular feature in the normal school routine. In long experiments with programs and machines the situation changes. Lane reports from the Jönköping experiment: "Attitudes towards programmed teaching were maintained very close to the neutral mean in the autumn and somewhat under this in the spring" (Lane, 1963, p. 24). In a study of the development of attitudes in students taking a programmed course in grammar, the attitudes of the students towards machine teaching was shown to become gradually less positive (Stukát, 1963).

1.1.3 *Examples of experiments to measure classroom behavior*

Few research projects include the task of measuring classroom behavior by means of systematic observations. Nevertheless it must be essential to map the interaction between teacher and students and the behavior of the students in the learning situation. In most learning experiments it is considered sufficient to measure previous knowledge and then note the effect of a special course. There are various underlying reasons for the fact that systematic experiments with classroom observers have hitherto been so rare: some teachers will not admit observers to their classrooms; it is difficult to find trained observers, which is essential to obtain reliable results; the observations are time-consuming and expensive to implement. The presence of observers in the classroom is a factor that, according to the teachers, affects the students. Noted behavior can perhaps not be regarded as typical of student behavior when the observer is not present.

A detail survey of experiments in measuring classroom behavior has been given by Medley & Mitzel (1964). Some of the information given there about studies conducted during the period 1950—60 will be presented briefly.

During this period attempts have been made to describe classroom behavior quantitatively. Classroom behavior is not as before placed in relation to teacher efficiency or psychological theories.

Cornell, Lindvall & Saupe (1952) measure *differences in behavior in order to characterize different teaching systems*. The following elements are noted in an extensive schedule: opportunities for differentiation, social organization, degree of student initiative, syllabus, occurrence of educational aids, teacher behavior, classroom atmosphere.

Medley & Mitzel (1958) developed from this an instrument called OScaR (Observation Schedule and Record). They wanted to increase the reliability of the observations and therefore simplified the schedule as much as possible. The observer was only to note what happened and not evaluate the behavior. With this schedule 49 teacher units were studied for 10 weeks. When the data was analyzed the activities that had been noted were grouped in categories.

The categories that showed reliable differences between different teaching units were factor analyzed, after which the categories were grouped in three dimensions according to the first three factors in the analysis. These factors were called Emotional Climate, Verbal Emphasis and Social Organization. The last two correspond partly to the categories working activity and goal-centered interactions in my observation schedule (Holmberg, 1965, appendix 2).

Wright (1959, 1961) studied teacher/student interaction during mathematics lessons, in order to map the verbal behavior. Her contribution to the methodology is a time-sampling system, which means that the lesson is divided up

into intervals of 15 seconds. The observer watches the verbal behavior for 15 seconds and classifies it for 15 seconds. Her observation schedule is complicated and makes great demands on the observer.

Many different observation schedules have been used. The observation technique has been improved, analysis methods have been developed—but it still remains to map the relative importance of the separate behavioral categories we noted in a learning environment. Each special study must in addition concentrate on its objectives during the observations. Despite the large number of existing methods, it has therefore seemed essential in this investigation to try a new observation schedule.

1.2 Problems

The problems we have set out to study in these intensive investigations, each of which extended over three terms, have been the following:

1. To investigate whether differences in achievements in mathematics can be shown between classes with one of the strategies of programmed instruction outlined above, (PI 1) and (PI 2) respectively, and classes with conventional instruction (CI).
2. To investigate possible differences in reading ability between classes with PI 1 and PI 2 respectively and classes with CI.
3. To investigate differences in observed behavior (working activity, play and passivity, goal-centered interactions and disturbing interactions) between classes with PI 1 and PI 2 respectively and classes with CI.
4. To investigate if and in what respect the attitude towards mathematics changes during a longer period in classes with PI 1, PI 2 and CI.
5. To investigate whether the preference for mathematics changes during a longer period in classes with PI 1, PI 2 and CI.
6. To investigate whether the attitude towards machine teaching changes in the machine classes PI 1 and PI 2.
7. To investigate possible differences in working habits between classes with PI 2 and classes with CI (applies only to Investigation of Effects 2).

1.3 Methods

1.3.1 *Experimental design*

The machine classes and comparison classes were studied continuously with regard to achievement, attitude and behavior variables during mathematics lessons for three consecutive terms. The machine classes used different methods

in the two experiments (PI 1 and PI 2). The comparison classes used the same methods in both experiments.

Achievements in mathematics were measured with test papers, which were given once a month.

Reading ability was investigated once a term. In the Investigation of Effects 1, Diagnostic Reading and Writing Tests (Skandinaviska Testförlaget) were used and in the Investigation of Effects 2, Reading Tests A and B (Lövgren, 1964) were used.

Behavior was registered in Investigation of Effects 1 for 24 students, 12 from each class, and in Investigation of Effects 2 for 36 students, 18 from each class. The observations concerning working activity, passivity and disturbing activity were noted for these selected, matched students. The behavior was registered in a schedule for observation of individual students (Holmberg, 1966, appendix 2).

In each case control observers were employed in order to make it possible to determine the reliability of the observations. A special classroom at the School of Education equipped with one-way screens permitted discreet observations of the classroom situation and discreet tape-recording of the classroom interaction. The observation technique is described in detail in Holmberg, 1965, p. 11.

The same observation methods were used for both experiments. These observations were made twice a month by student teachers at the School of Education, who had been specially trained in observation technique.

Preferences for the subject of mathematics were tested twice a term with Form H, an instrument for measuring schoolchildren's preferences in school subjects (Holmberg, 1965, appendix 1 and Holmberg, 1966, appendix 1).

Data on attitude changes towards machine teaching were gathered in Investigation of Effects 1, by letting the students write an essay with the title: "What I think about being taught by machine." In this essay they were to answer six questions which were given in the instructions for the writing of the essay. Students participating in Investigation of Effects 2 answered Attitude Test H, an instrument developed from the instructions for essay writing that had been used in the first experiment (Holmberg, 1966, appendix 3).

Question-frequency notation from chemistry experiments and *instruction-comprehension tests* have been aids in the investigation of working habits (Holmberg, 1966, appendix 4). The investigation of working habits applies only to Investigation of Effects 2.

1.3.2 *Subjects of experiments*

The subjects in Investigations of Effects 1 and 2 consisted of two classes who in the autumn terms of 1962 and 1963 respectively started the 7th grade at

Heleneholm School in Malmö, Sweden, so-called occasional samples. Practical consideration affected the selection of these subjects for long-term experiments. It was necessary to have access to two classes with the same teacher. Proximity to suitable observation rooms was also a prerequisite for the implementation of the experiment. By means of careful observations it would be possible to map the learning situation and its development in the two classes with different methods and describe the development. In order to have a foundation for making generalizations, the representativeness of the class was checked.

For the students participating in Investigation of Effects 1, the following was found:

Intellectually the experimental group is representative of the general course in the 7th grade of the compulsory school. F test (the F test in Form A, 1960): $m = 4.0$ (stanine points). Totally for 200 students from the general course the figures are $m = 4.0$ (Öreberg, 1964).

The requirement of intellectual comparability between the classes has been observed. In testing with Cattell Culture Fair 2A (Cattell, R. B. & Cattell, A. K. S., 1965) the experiment class and the comparison class both had $m = 25$. The spread was 5.77 and 6.33 respectively. In order to be able to compare the students' previous knowledge of mathematics, a test paper was given. No significant difference between the classes could be established ($p > 0.20$, $\omega^2 = 0.02$).

Regarding the behavior variables, more detailed observation data has not been gathered for the whole classes. By means of matching, 12 students were chosen from each class for special observation. The students were selected so as to be comparable intellectually (Cattell points) and in mathematics (test paper results).

From each class high group students, average students and low group students were chosen, with reference to achievement and intelligence.

For the students participating in Investigation of Effects 2 the following was found:

The two experimental groups are intellectually comparable. When tested with Cattell Culture Fair 2A (Cattell, R. B. & Cattell, A. K. S., 1965) no significant difference was found ($p > 0.20$; $\omega^2 = 0.03$).

Previous knowledge in mathematics was investigated by means of a test paper, in which both classes achieved similar results ($p > 0.20$; $\omega^2 = 0.02$).

Observation data was also collected on all students. At the end of the experimental period we matched 18 students from each class. The experimental groups were divided into high, average and low groups according to the results from the pre-tests (mathematics test and intelligence tests, Cattell Culture Fair 2A).

1.3.3 *The role of the teacher*

In Investigation of Effects 1:

The comparison class and the experimental class had the same teacher. Since we wished to study the reactions of the students to the programmed material, the teacher was instructed to remain as passive as possible in the machine class. The teacher was forced to intervene sometimes, however, to explain further to weak students in some cases and to adjust small faults in the teaching machines during the autumn term of 1962.

In Investigation of Effects 2:

The same teacher as in the previous experiment has taught both the machine class and the comparison class. In the machine class his task has been to complement the machine teaching with small-group teaching. The results from the previous experiment with maximum use of teaching machines during a longer period (Holmberg, 1965) had namely shown that the attitudes of the students towards programmed teaching had become significantly more negative. For this reason it was considered essential to modify the role of the teacher in this new experiment. In the first investigation the teacher was instructed to remain passive in the machine class. In this experiment the teacher was to motivate the students by alternating individual work with teacher-led small-group teaching. The teacher was asked to adapt his teaching to the needs of different groups of students. Different program material was to be chosen by the teacher and student jointly.

1.3.4 *Teaching methods*

Investigation of Effects 1:

One class was given machine instruction, while the parallel class had teacher instruction. The comparison class was taught by means of a kind of traditional class teaching. The teacher explained new items on the board. The students practised the calculation: first on the board and then individually in their books (Textbook: Wahlström, B. & Olsén, A. Matematik. Grundskolans årskurs 7. Allmän kurs. Stockholm: Svenska Bokförlaget, 1962). The students could check their answers in a key. The teacher helped individual students when the class was working independently. All the students followed the same sequence in the material. A certain amount of individualization occurred since some slow students were given a reduced number of calculations to do and only those who wished to did the extra tasks.

For the machine teaching we used the Swedish PIL apparatus, which is

based on Skinner's principles for learning. The letters PIL stand for the Swedish, *Programmerad InLärning* (English = Programmed learning). The program, the only one available in Swedish, was a linear program of the Skinner type (Jansson, 1963, stencil). The machine and program are described in detail in Holmberg, 1965, p. 14.

No homework was given in the machine class. The comparison class had homework once a week.

Investigation of Effects 2:

One class has had modified machine teaching, while the parallel class has had the same textbook and the same class teaching as the comparison class in Investigation of Effects 1.

For the machine teaching we have as in the previous experiment used the PIL apparatus and a revised edition of the program used in Investigation of Effects 1 (Jansson, 1963). Students who worked at approximately the same speed with the programmed material stopped working with the machines once or twice a week and under the guidance of the teacher did mental arithmetic and mathematical problems appropriate to the stage they had reached in the course. The time for the group teaching could vary between 10 and 20 minutes. The size of the groups varied between 2 and 6 students. There was also individual tutoring. During the last term of the experiment, three groups crystallized, the members of which could keep up roughly the same rate of work.

1.3.5 Measuring instruments

1.3.5.1 Mathematics tests

The test papers have provided the main measurement of the effectiveness of the mathematics teaching and have been used to compare the two varieties of the programmed course with conventional teaching. The various parts of the mathematics course are presented in the same sequence in both the textbook and the program. Despite this, it proved to be very difficult to put together test papers, which suited both classes, since at any given moment the students were all at different points in the course. The spread was very large in the machine class. While the slow ones did not manage more than a couple of series, the quick workers had completed ten or more during the first month of the experiment. For this reason the test papers were given to the machine class individually after certain series. In the comparison class the students were all given the test papers at the same time.

The reliability of the tests has been determined by means of the split-half

method and varies between $r = 0.95$ and 0.56 . Most of the test papers have been put together by the class teacher. The tests for April and May 1965 are included in the IMU project (Öreberg, 1965).

The standard tests (for the whole of Sweden) in mathematics for grade 8 are reported. The results provide a measurement of the general standard of the experimental classes.

1.3.5.2 Reading test

Investigation of Effects 1:

At the start of the project no standardized test in reading for grade 7 existed. In order to obtain data on the development of the reading ability in the classes concerned and to be able to compare the reading achievements of the two experimental classes, we used DLS (Diagnostiska Läs- och Skrivprov, 1964; reading comprehension and reading speed tests for grades 4—6). No standardization data for the test could be used. The general standard of the reading ability in the two classes, i.e. the average level in relation to the mean for the whole country, is reported by means of the results of the standard tests in grade 8.

Investigation of Effects 2:

Reading ability has been investigated by means of the then newly-constructed reading tests A and B (Lövgren, 1964). The tests have been used to obtain a comparison between the reading achievements of the two experimental classes and to get data on the development of reading ability in the classes. In addition the tests provided information about the average level of the classes in relation to other classes in grades 7 and 8 of the comprehensive school.

1.3.5.3 Observation schedule

Investigation of Effects 1:

After preliminary studies of the survey group and other parallell classes in grade 7, a schedule was worked out for the observation of individual students. The activities included were those that could be considered to cover the course of events during a mathematics lesson, to be connected with the formulation of problems. The observation schedule contains 26 activities. In order to make the schedule easy to read and prevent incorrect registrations, these activities were divided into categories: introductory and terminating activities, individual work, play and passivity, interactions. When the material was processed, the

interactions were divided into disturbing and goal-directed interactions.

The reliability of the observation schedule is described in percentages of the agreement between observers, see Holmberg, 1965, p. 18.

It proved to be very difficult to mark behavior within individual variables. For the variable "Individual work," the percentage of agreement varies between 84 % and 53 % for the machine class and between 82 % and 44 % in the comparison class. Thus the percentage of agreement is low. It should be noted, however, that some percentages are based on very few markings and so are of little value as reliability data. In most comparisons the agreement between observers is better for markings concerning the machine class. It is probably simpler to note behavior of students working with machines.

Investigation of Effects 2:

The schedule that has been used in observing individual students is a revised version of the schedule that was used in the previous experiment (Holmberg, 1966, appendix 1). The schedule was revised since we wished to change some working routines, which resulted in there being fewer activities to register. In addition it emerged that in some categories of behavior we should combine the markings in order to obtain a more reliable schedule. The activities "Helps classmate with calculation" and "Gets help with calculation from classmate" become, for example in the revised version "Gets help or helps classmate with calculation." The new version of the observation schedule contains 19 activities, divided between the same number of categories as before.

The reliability of the observation schedule is reported in percentages of the agreement between observers (Holmberg, 1966, p. 11 ff.). For the variable "Individual work," the percentage of agreement is 92 % (m) for the machine class and 81 % (m) for the comparison class. It should be noted, however, in this investigation too, that in variables with few markings e.g. "Goal-directed interactions," the percentages have little value as data of reliability. In this experiment also it was found that the agreement between observers is better for markings concerning the machine class. It is easier to note the behavior of students working with machines than that of students who are expected to listen to the teacher's explanations of the subject.

1.3.5.4 Form H: an instrument for measurement of subject preferences of schoolchildren

Investigation of Effects 1:

In order to find out the extent to which a change of teaching method in mathematics influences the preference for this subject, a form was constructed

with paired comparisons (Holmberg, 1965, appendix 1). This form takes up school activities, which occur in grade 6. The information about the various subject areas has on the whole the same frequency as the subjects themselves have on the timetable. The form is answered in such a way that the subject with which the student would most like to work is marked. The number of choices per subject is added up, then the subjects are ranked.

We have tried to establish the validity of Form H by means of a logical analysis from empirical data (cf. Kerlinger, 1964, p. 45). We know that for students at higher levels moderate positive relations have been shown between achievements in mathematics and attitudes towards the subject (Alpert et al., 1963, Werdelin, 1966, Degnan, 1967). In the international mathematics investigation (Husén, 1967), it was noted that the relation between interest in mathematics and achievement is high or moderate. Considering these results, we have chosen to report extreme-group validity for the preference schedule. The difference with regard to mathematics preference points has been tested for the 27 % of the students with the best grades and the 27 % with the worst grades. The best students in mathematics here had significantly higher preference points than the worst ones (Mann-Whitney's U Test; Siegel, 1956, p. 116 ff.), $p < 0.01$.

The reliability is reported by means of ranking correlation-coefficients for re-testing: Sept. '62—Nov. '62 comparison class $r = 0.82$. Sept. '62—Nov. '62 machine class $r = 0.77$.

Investigation of Effects 2:

In this investigation Version III of Form H has been used (Holmberg, 1966, appendix 2).

1.3.5.5 Essay writing—Attitude Test H

Investigation of Effects 1:

The students in the machine class have written essays about their experiences of teaching with programmed material in Dec. 1963 and Dec. 1964. They were given the following stenciled instructions:

"Write an essay with the title:

What I think about being taught by machine.

Tell:

1. what you think of the machine
2. whether you think that you learn more or less by machine

3. whether it was more fun with machine teaching than with normal lessons with a teacher
4. whether you got more tired during a lesson with machine teaching
5. whether time goes more quickly
6. whether you would like to have machine teaching in more subjects."

Investigation of Effects 2:

When measuring attitudes towards programmed material in this investigation, we have developed from the instructions given for the essay-writing in the previous experiment a Likert Scale—Attitude Test H. (Holmberg, 1966, appendix 3.) For each statement the student is to mark his attitude on a numerical scale with five categories.

1.3.5.6 Question-frequency notation and instruction-comprehension test for measuring working habits (applies only to Investigation of Effects 2)

In order to obtain further information about whether the PI method of work changes the working habits of the students, special observations were made during chemistry lessons. After the first experiment with programmed teaching (Holmberg, 1965) the teacher in laboratory subjects noticed that the students in the machine class read instructions better than the students in the comparison class. The students in the machine class appeared to get to work more quickly and they asked the teacher for advice more seldom. It was therefore important to obtain data with which to check whether the machine class students really did ask the teacher for advice less frequently.

During the autumn term of 1965 an observer noted on the class placing schedule the question frequency for the two experimental classes during every other chemistry lesson. The observer was given careful instructions as to which questions were to be noted. The students were working with written instructions (Mårtensson, Strelert & Wernfors, *Laborationer i kemi*. Lund: Gleerup, 1964).

At the beginning of each lesson the chemistry teacher went quickly through the laboratory experiment and distributed material. The teacher then remained relatively passive while the students worked on their experiments. During this period (the students' independent laboratory work) the observer noted the frequency of questions concerning the laboratory work. The purpose has been to find out the extent to which the students managed the experiment solely with the help of their textbook. At the end of the lesson the teacher controlled the results of the experiment with the students and during this period no questions were noted.

Once a month a second observer made notes, so that we should be able to determine the reliability of the observations. The reliability is described in percentages of the agreement between the observers; this varies between 100 % and 91 %.

The ability of the students to read instructions has been measured by 10 items from Werdelin's instruction-comprehension test (Werdelin, 1958, p. 297). This test (Holmberg, 1966, appendix 4) has been given in September 1964 and September 1965.

1.3.6 *Processing methods*

When the first investigation was planned, the intention was only to analyze the difference between matching pairs. Knowledge of mathematics and intelligence were the matching variables. When data had been collected on all the variables except the behavior variables, however, the analysis in Investigation of Effects 1 had been complemented by a comparison of the undivided classes (t analysis). Further the results of the students in the high, average and low groups has been analyzed for variance, since the results proved to be difficult to interpret by means of the methods we had originally planned to use. The model that has been used is related to an example by Lindquist (1956, p. 281 ff.). The missing values have been corrected by Lindquist's method (Lindquist, 1956, p. 148 ff.).

The same analysis of variance model has been used in Investigation of Effects 2. The analyses of variance in Investigation of Effects 2 have been complemented by some t analyses and comparisons between matching pairs. Certain correlation calculations have been carried out in both investigations.

The level of significance has been placed at 0.05. The effect size in the tests has been estimated by means of Hays's ω^2 (Hays, 1963, p. 325 and 406). We have used ω^2 0.05 as the criterion for essential differences both for the analyses of variance and the t tests (cf. Cohen, 1969, p. 277 and 359).

1.4 Results and discussion

1.4.1 *Analysis of changes in achievement in mathematics*

In the analyses of variance it could be established that the teaching methods in both experiments produced equivalent results when taken as a whole ($p > 0.05$, $\omega^2 = 0.00$).

There were differences between the high, middle and low groups. The high group achieved the best results and the low group the worst. This is natural, since the division into groups was based on achievements in mathematics ($p < 0.05$, $\omega^2 = 0.07$ resp. 0.03).

In the first experiment the two methods had a different time-pattern: the machine class was ahead of the comparison class at the beginning of the experiment (autumn term '62) but behind at the end of the experiment (autumn term '63) ($p < 0.0005$, $\omega^2 = 0.73$).

In the analysis with t tests of the results for unstreamed classes, Investigation of Effects 1 established that there were no significant differences in the results of the test papers between the two classes on any of the test occasions during the three terms the experiment lasted.

In a comparison with a national sample for Investigation of Effects 1, no significant difference could be noted between either of the experiment classes and this national sample (standard test in mathematics for grade 8, spring term '74). The differences in achievements in mathematics were tested by means of t tests ($p > 0.25$, $\omega^2 = 0.00$ for both classes). Thus the experiment classes can be considered representative of their age-group in Sweden.

The machine class was significantly better than the comparison class in the standard tests ($p < 0.05$, $\omega^2 = 0.07$). After having used machines and mathematics programs for three terms, the attitude of the students towards this form of teaching was negative, but in their achievements they held their own.

1.4.2 *Analysis of changes in achievements in reading ability*

Different types of tests have been used in the two experiments and therefore the results are reported separately.

Investigation of Effects 1:

In the analysis of variance no differences were noted in either reading speed or in reading comprehension between the two methods compared ($p > 0.05$, $\omega^2 = 0.06$). Differences in reading speed were found between the groups. The high group read more quickly than the others. In reading comprehension no difference could be noted between the groups.

The interaction "method x test occasion" showed that the time-pattern was different for the two methods ($p < 0.05$, $\omega^2 = 0.05$). In the machine class reading comprehension increased steadily, while reading speed fluctuated from occasion to occasion.

In the analysis by means of t tests of the results for the whole class, no significant differences emerged in either reading speed or reading comprehension during any one term of the experiment ($p > 0.10$ on each testing occasion and ω^2 varies between 0.01 and 0.00).

When *compared with the national sample*, no significant differences were found to exist between the groups compared (t tests on the standard tests for grade 8, the test reading comprehension: $p > 0.25$, $\omega^2 = 0.00$ for all groups).

The students in the machine class did admittedly have the highest mean in reading comprehension, but if that was an effect of PI, the effect was not large enough to produce any statistical evidence of the possible superiority of the method. The general standard of the experiment groups in reading ability did not in other words deviate from the mean for the whole of Sweden.

Investigation of Effects 2:

In the *analysis of variance* no significant differences in reading ability were found between the methods compared ($p > 0.05$, $\omega^2 = 0.00$). Nor did the group division show any difference in reading ability. The difference between the test occasions was significant ($p < 0.001$, $\omega^2 = 0.15$)—reading ability improved continuously and this trend could be shown most clearly in the comparison class (Holmberg, 1966, appendix 5: 3).

In a comparison with a Stockholm sample (standardizing data) our experiment groups proved to be comparable with these upper level students in reading ability. No significant differences occurred between the experiment groups and the Stockholm sample ($p > 0.20$ for both groups, $\omega^2 = 0.03$ and 0.00 respectively).

1.4.3 *Analysis of behavior assessments*

Investigation of Effects 1:

In the analyses of variance no significant differences in behavior can be shown between the students in the machine class and the others. Analyses have been made on the variables “Individual work,” “Play and passiveness,” “Goal-directed interactions” and “Disturbing interactions.” The group division produces the only significant difference in the variable “Disturbing interactions” ($p < 0.005$, $\omega^2 = 0.04$). The middle group causes more disturbance than the rest of the students and this is particularly marked towards the end of the experiment period (Holmberg, 1965, Fig. 11, appendix 13). The variable “Play and passiveness” has the only significant interaction “group x test occasion” ($p < 0.001$, $\omega^2 = 0.13$). In the high group the score is lowest during the first term, in the middle group it is relatively similar throughout the three terms, while the low group has maximum score during the second term. This is the only variable that shows varying time-patterns for low, middle and high groups. The high groups started in the upper level of the comprehensive school by working intensively during the mathematics lessons—they had no time to play. By the second term the situation had changed. The students in the

machine class thought that the programmed material was too easy and consequently boring, so the work activity diminished. The high group students in the comparison class had to wait for their classmates to catch up during the teacher-led parts of the lesson and the waiting time was used for playing. The time-pattern in the behavior variables shows differences in the two experiment classes, which are the result of irregularity in the time-pattern rather than of a distinct trend. As far as levels are concerned, there are no differences between methods, as can be seen in the curves in Figs. 8—11 (Holmberg, 1965, app. 10—13), in which the curves of the two classes either run parallel or cross each other. In summing up it can be said that no differences that are relevant to the problems taken up in the investigation exist.

Investigation of Effects 2:

In the analyses of variance, behavior during mathematics lesson is shown to be rather similar in both classes. There is a significant difference in level between the methods only in the variable “disturbing interactions” ($p < 0.001$, $\omega^2 = 0.15$). The students in the comparison class cause most disturbance. The disturbing activities are limited, however, to a small proportion of the activities that occur; 70—80 % of the time is devoted to individual work, which must be said to reflect a good working environment.

The group division produces no significant variation in any variable.

In all the behavior variables we find a significant variation between the test occasions. The high variation between the test occasions ($p < 0.01$, ω^2 varierar mellan 0.10 och 0.13), is a result of irregularities in the time-pattern rather than of a distinct trend.

The variable “individual work” has significant interaction “method x test occasion” ($p < 0.01$, $\omega^2 = 0.07$) which means that the time-pattern for the two classes was different. During the first term the markings for individual work increase greatly in the machine class and during the second term the markings decrease, only to increase again in the third term. In the comparison class the number of markings decreases rather a lot for individual work during the third term.

The variable “disturbing interactions” also shows significant variation, “method x test occasion” ($p < 0.001$, $\omega^2 = 0.07$). In the comparison class the disturbing activities increase at the end of each term, reaching a climax at the end of the third term. There is no corresponding pattern in the machine class. Obviously the modified model with alternating group and machine teaching provides much better prospects for a good work environment, which encourages fewer disturbing activities than the conventional method.

1.4.4 Analysis of preference for mathematics (Form H)

Investigation of Effects 1:

The analysis of variance established that neither teaching method nor group division appeared to have any importance for the mean preference over the entire experiment period ($p > 0.05$, $\omega^2 = 0.01$ resp. 0.04).

The preference for mathematics showed different time-patterns within the two teaching methods (Holmberg, 1965, app. 14, Fig. 12). The course of development was parallel to the development in achievement, i.e. the machine class lay at the top in the beginning, but the preference score in the machine class sank steadily throughout the entire experiment. The groups also had different time-patterns and this development is responsible for the greater part of the significant variation between the test occasions.

In the machine class the following points were noted: In the beginning there was great interest in the new method among the best students. This interest then diminished most rapidly within the *high group*, where the students first pointed out the deficiencies of the programmed material. Their complaint was that the frames were so simple that they lost interest, the repetition came too quickly and was too abundant. The students found the programs boring.

In the *middle group* the interest in PI lasted longer. The decline in preference scores did not come until the Spring term 1963. During the autumn term '63 the dissatisfaction with PI was accompanied by deterioration in the results on the test papers. In the test papers for both October and December '63 the comparison class achieved significantly better results.

The *low group* had initial difficulties, which are reflected in the results on Form H (Holmberg, 1965, app. 15, Fig. 13). During the spring term '63 the preference for mathematics increased somewhat, diminishing again substantially during the autumn term '63.

In the comparison class there was also an increase in preference scores at the beginning of the autumn term in grade 7. The same effect was noted when Form H was used in the IMU project (Swedish Board of Education's investigation into individualized mathematics teaching in grade 7; Öreberg 1964). After this initial increase the curve for the whole class remains relatively horizontal. The preference scores of the *high group* show the greatest variation between the different test occasions. On all test occasions they are higher than in the high group of the machine class. The mathematics preference of the *middle group* produces an even curve. It is lower than that of the machine class during the first two terms of the experiment, but in the autumn term '63 the curves cross each other.

Investigation of Effects 2:

In the *analysis of variance* it was noted that neither teaching method nor group division had any significant influence on the mean preference for the experiment period as a whole.

The variation between the test occasions was significant, as it was in the former experiment ($p < 0.005$, $\omega^2 = 0.06$). Comparison of the development of mathematics preference in these two experiments shows that lowering of the preference below initial level, which occurred in the first experiment in the machine class, was completely absent in the second experiment. In experiment 2 we again had an initial increase in the machine class, which undoubtedly resulted from the enthusiasm felt for this new teaching method. The initial enthusiasm produced the greatest increase in preference scores for the high group. The students in the high group also expressed their approval of the variation in the teaching method: group work alternating with programmed instruction.

1.4.5 The attitude towards machine teaching as reflected in essay-writing or measured by Attitude test H (applies to students with machine teaching)

Investigation of Effects 1:

In the first experiment the students were asked to write essays on their experience of teaching machines. The essays have then been analyzed from the point of view of the students' reactions to six points that were stated in the instructions for the essay-writing. The answers on each point have been assessed according to a five-point scale: + +, +, 0, -, - -. The scale has been scored (1—5) with five points for the most positive answer and one point for the most negative.

The essays have been analyzed by two independent assessors. The assessor reliability for grade 7: $r = 0.98$; for grade 8: $r = 0.99$.

The means for these two assessments have then formed the foundation for comparisons between the essays written in grade 7 and those written in grade 8. The student reactions in grade 7 are rather positive for the majority of the students. This positive attitude in students taking the general course in mathematics in the comprehensive school is remarkable when compared to the results of other attitude studies (e.g. Johannesson, 1960, p. 80). Here is a short essay that is typical of grade 7:

"What I think about being taught by machine.

I have been doing maths by machine for nearly a year now and I think it is good, although it breaks down sometimes. It is oblong and looks like

a rectangle. Counting with it is quick and more fun and I think I have learned more than with an arithmetic book. You don't get so tired either and time goes more quickly. I'd like to have a machine in more subjects."

After a further year's machine teaching, the same student wrote:

"What I think about being taught by machine.

It is very dull and useless!

I think I learned much less sitting there and doing sum after sum in the same way. Time went more slowly and you got so slack just sitting and turning a screw; no, down with these machines. And I wouldn't want to have a machine in more subjects for anything in the world."

The difference between the results for the entire classes has been measured by means of *t tests*. The students are significantly more negative in class 8 than in class 7 ($p < 0.01$, $\omega^2 = 0.20$). This result must be seen in relation to the decrease in interest that is normal from grade 7 to grade 9, irrespective of which subject is investigated (cf. Flinck, 1965, p. 87). This change in attitudes also seems to be independent of, in which decennium the investigation is carried out. Rudvall reports from the VGT experiments (Varying group-size and teacher teams): "The general tendency is a positive or neutral attitude towards school work in grade 7 . . . On the whole the interest in school work diminishes somewhat for each new report in grades 7 and 8, then stabilizing on the grade 8 level in grade 9." (Gran, B., Gran, Bi., Rudvall, G. & Röhr, G., 1974, p. 55).

Investigation of Effects 2:

In the second experiment Attitude test H was used to evaluate the attitudes of the students towards machine teaching. This instrument had been developed from the instructions for the essay-writing. The students in the machine class answered Attitude test H in Dec. '64 (grade 7) and Dec. '65 (grade 8). If the scores on all the statements in the test are added together and the difference in attitude between grades 7 and 8 measured (*t test*), a significant negative change can be noted ($p < 0.05$, $\omega^2 = 0.12$) even in this modified experiment. When the answers to the items are studied, a more differentiated picture of the attitude development emerges.

In grade 7 over half the students think that programmed teaching is as much fun or more fun than conventional teaching. In grade 8 the situation is the reverse. This change in attitude is not so great, however, that the difference from grade 7 to grade 8 between the results for the entire grades are significant ($p > 0.20$, $\omega^2 = 0.02$).

After three terms' teaching with programmed material, 67 % of the students consider that they learn more or as much mathematics as they would have done with conventional teaching.

After three terms' work 73 % of the students find programmed teaching more tiring, since they must concentrate more and obtain almost all information themselves by reading.

Somewhat more than half of the students reply negatively to the questions whether they want machine teaching in revision of the mathematics course.

In grade 7, 14 % of the students do not want any machine teaching at all, while 9 % want only machine teaching. The majority prefer variation in the type of teaching.

In grade 8, 47 % of the students state that they do not want any machine teaching at all, while 7 % want only machine teaching. The rest of the students desire alternation between the two types of teaching. The change in attitude from the one year to the next is not so great as to make the difference significant ($p > 0.20$, $\omega^2 = 0.01$).

1.4.6 *Analysis of working habits (applies only to Investigation of Effects 2)*

The *analysis of ability to read instructions* (tested by means of Werdelin's instructions comprehension test) showed that the machine class significantly increased its ability ($p < 0.005$, $\omega^2 = 0.05$) from grade 7 to grade 8. In this the constant work with programmed material can be a contributory cause. In the comparison class the mean of the students' scores on the instructions comprehension test increases less, however, and the difference between grades 7 and 8 is not significant ($p > 0.01$, $\omega^2 = 0.01$).

When the difference between the machine and the comparison classes is tested, the result is:

Sep. '64 $p > 0.20$, $\omega^2 = 0.02$

Sep. '65 $p < 0.05$, $\omega^2 = 0.06$

At the start of the experiment period there was no difference between the ability of the two classes to read instructions. After two terms the machine class was significantly superior. The machine class has improved more than the comparison class. When the difference in the increase in test scores is tested, the result is $p < 0.01$, $\omega^2 = 0.14$. This result confirms the subjective assessment of the teachers that students who have worked with programmed material are better at reading instructions and carrying out laboratory experiments independently. This development is also in complete agreement with the broader goal of the school to give the students good working habits and a suitable study technique. "The school should teach students . . . to acquire skills

and knowledge independently and methodically . . ." (Curriculum for the comprehensive school, 1962, p. 16.)

During *observation of the students' working habits* during chemistry lessons it was noted that the students in the machine class tended to be more independent and that more questions were asked in the comparison class.

The results of the observations made during chemistry experiments have been analyzed for variance according to the model described earlier. Taking the experiment period as a whole, the analysis of variance gives no significant difference between the methods ($p > 0.05$, $\omega^2 = 0.05$).

1.4.7 *Some studies of relations*

1.4.7.1 Mathematics achievement—preference for mathematics

In both investigations of effect weak relations are reported for all the classes between mathematics achievement and preference for the subject (r varies between 0.02 and 0.40). The group with best achievements has significantly higher scores than the worst group. In my experiment groups the middle groups have unusually high preference scores. This indicates a non-linear relation, which can explain the low coefficients which refer to linear relations.

1.4.7.2 Mathematics achievement—reading ability

Investigation of Effects 1:

Reading speed (RS) and reading comprehension (RC) have been correlated with achievements in mathematics—number of correct answers on test papers (M). Partial correlation coefficients have been worked out between reading speed and mathematics achievement and between reading comprehension and mathematics achievement. A multiple correlation coefficient between mathematics achievement and reading speed + reading comprehension is also given (Holmberg, 1965, p. 41).

In the machine class a high reading speed—without accompanying good reading comprehension—appears to lead to worse results, while in the comparison class reading speed has been of no importance. The latter fact is only to be expected and it is also natural if a quick perusal of the tasks without much understanding of the content leads to bad results, since there is a time-limit in operation. It is equally understandable that increased reading comprehension leads to better results. One peculiar circumstance is that while the relation between reading speed and reading comprehension is rather high in the machine class in the autumn terms of 1962 and 1963, there is no such relation for the spring term 1963.

Investigation of Effects 2:

In this investigation only one value is given for reading ability: a reading test score. The relations between reading test score and mathematics achievement—the number of R on the test papers—are rather low. The weakest relation was noted in September '65, $r = 0.21$; the highest relation was noted in May '65, $r = 0.56$.

1.4.7.3 Correct answers in the test papers—correct answers in the series in the programmed material (applies to students in the machine classes)

All students had not done the same number of series at the end of the term. The spread was rather large. Some quick-working students were allowed to start on the course for the next grade, while some slow students were allowed to by-pass some revision series. The percentage of errors for the series varied between 4.8 % and 15 %. The relation between the percentage of correct answers on the series done and the percentage of correct answers on the test papers has been calculated.

autumn term	'62	$r = 0.37$
spring term	'63	$r = -0.47$
autumn term	'63	$r = -0.42$
autumn term	'64	$r = 0.62$
spring term	'65	$r = 0.38$
autumn term	'65	$r = 0.30$

Weak positive and negative relations both exist. These results are in agreement with earlier research results (Gagné & Dick, 1962). The programmed material is probably made far too easy with its numerous clues, which gives the students too little training in problem-solving. The test situation presumably makes demands on the student that the programmed material has not prepared him for.

1.5 Summary

1. In the *achievement variable mathematics*, the different methods produced equivalent results in both experiments.

In the standard tests for grade 8 there was in the spring term of '64 no significant difference between any of the experiment classes and the national population. On the other hand the students of the machine class with a maximum use of teaching machines were significantly better than those of the comparison class.

2. In *reading ability* no differences occurred between the classes compared if we consider the experiment period as a whole.

According to the standard tests and standardized reading tests the experiment groups are representative of upper level students in general.

Reading comprehension increases steadily for students with a maximum use of teaching machines, while reading speed fluctuates from occasion to occasion.

3. *Behavior*: More disturbances occurred in classes with conventional teaching than in classes with machine teaching combined with group teaching. At the end of each term the disturbing activities increased in classes with conventional teaching.

Behavioral differences between high, middle and low groups were only noted in experiment 1 and then in the variable "Disturbing interactions." The middle group here had higher scores than the others, especially towards the end of the experiment period in grade 8. Disturbances encroached on very little of the lesson time, however.

The great variation between the test occasions for all behavior variables depend on irregularities in the time-pattern, rather than on a distinct trend.

4. *Preference for mathematics*, measured by Form H, remained relatively constant for the comparison classes. In the class with maximum use of teaching machines the preference diminished, but there was still no significant difference between the machine and comparison classes. In the class with the modified experiment (group teaching + machine teaching) the preference values for the whole experiment period lay somewhat above those for the comparison class, but no difference between the methods could be established.

5. *Attitude towards machine teaching* has been studied by means of essay analysis in Experiment 1. The reaction of the students in class 7 was fairly positive. This changed in grade 8. The students were significantly more negative in class 8 than in class 7. After three terms' teaching by machine 15 out of 16 students had a negative attitude towards PI. This result should not be over-interpreted. We have no information as to how negative the comparison class was towards its mathematics teaching in grade 8.

In Experiment 2 the attitudes towards machine teaching were evaluated by means of Attitude test H. When the scores on all the statements in the test were added and the difference in attitude between grade 7 and grade 8 was tested, a negative bias was noted in this modified experiment too. This difference stemmed mostly from the circumstance that the students considered programmed teaching more tiring than conventional teaching. After three months' work with machines, 73 % of the students found this form of work

more tiring. 67 % of the students thought that they learned more mathematics or as much as they would have done with conventional teaching.

Thus in our two long-term experiments—the first with maximum use of teaching machines and programmed material and the other with programmed material presented via machines alternating with teacher-led small-group teaching, both compared to conventional teaching—no significant differences were found in mathematics achievement or in preferences for the subject. Behavior during mathematics lessons is rather similar for students using the three different methods. The attitudes towards the method with maximum use of teaching machines is negative. When we modify the method by alternating group-teaching and machine-teaching, the attitudes become more moderate. The students consider that programmed teaching is more tiring than conventional teaching, but in compensation they think they learn more. This opinion is to certain extent verified by the results obtained on the standard tests, which were given three weeks after the completion of Investigation of Effects 1.

In the standard tests in mathematics, spring term '64, the class with a maximal use of teaching machines was significantly better than the comparison class. This is an important result, since the standard tests consist of centrally-produced test material. The good results achieved by the machine class in the standard tests show that the programmed material has given the students good training in various mathematical skills. The doubts expressed earlier that PI might make the students less able to solve problems were not confirmed. The students in the machine class achieved these good results in the standard tests despite the fact that they have had less time for mathematics training than those in the comparison class, who had homework once a week.

1.6 Conclusions

Further consideration must be taken to the tiredness factor. The teaching methods could be alternated more frequently than has been the case in our modified experiment. Perhaps short programmed courses on individual items and extra supplementary courses for the fast-working students would be a sensible way of using the program method. Programmed material with systematically applied teaching principles should also have a function to fill in remedial teaching, as a mean of individualizing the work.

The teacher should be a central and active component in mathematics teaching. The machines cannot fill the social function required by the students.

It is naturally difficult to generalize on the basis of a few experiments. Many separate factors can have influenced the results. In the comparison classes, for example, the students have been given homework, which has given them more time for practice. It is impossible to determine the effect of this extra

learning. On each test occasion the students in the machine classes had reached different points in the course. Under such circumstances it is difficult to make wholly fair comparisons. A lack of reliability in some test papers—to some extent because the tests were made too short—can also have affected the results.

In these experiments the effect of a program has been compared to the effect of instruction by an experienced and skillful teacher. Although the program was only at the trial stage, the classes had attained comparable achievements after three terms of experimentation. One may assume that the result would have been different if the program had been finally revised and if the teacher in the comparison class had been less used to teaching at this school level.

Indirectly the method of programmed teaching has stimulated the educational debate. Interest in the production of individualized teaching material has increased and greater demands are being made on the effectiveness of such material. In order to check the learning effect, trial productions have been intensified. "The strength of programmed material is that its effectiveness can be empirically improved and measured," (Bjerstedt, 1963, p. 236). Teachers have become more aware of the principles for the construction and evaluation of material and are now demanding more carefully written textbooks.

The teacher in the machine class has been forced into new working habits by the use of programmed material. He has been the co-planner, who has given individual tutoring, noting at the same time which students could be placed together in small groups for teacher-led group-instruction. He has also been able to distribute the programs according to the needs of individual students.

The programmed teaching has partly changed the students school situation. They have had to work actively with the material the whole time in order to obtain information. Most of the students have found this continual activity tiring. In the modified experiment the time allotted for programmed work appears to have been more than sufficient. Some students appear to be quite unsuited to this form of programmed teaching and they would have been more motivated in their work if they had been freed from the machines. It takes time to learn a new teaching method effectively. A livelier alternation between teacher-led instruction and individual work by the students, together with a more flexible adaptation of the method in relation to the types of student would perhaps produce an optimal learning climate.

Towards the end of the 1960's, the big mathematics projects in Sweden (IMU, the Lund project, Nordic Committee for the Modernization of Mathematics Teaching) were working on a number of aspects of mathematics as a school subject: the goals of mathematics teaching, its organization, content and new methods. The numerous experiments trying out isolated changes in

methods and organization had proved to produce little improvement. The emphasis was instead placed on the importance of linking an organization model with syllabuses and teaching procedure.

Experiences gained in this research contributed to the writing of the mathematics section of the new curriculum and the supplement Lgr 69 II: Ma with directives and comments.

Since the introduction of the curriculum in 1970, research is being done to evaluate the effects of the new line in mathematics. This research will form the basis for a continuous review of the curriculum. The next part of this thesis, Investigation of Effects 3, describes the effects of Lgr 69 on certain aspects of the development of the concept of number, knowledge of mathematics and attitudes towards the subject.

2. Comparison of curricula: investigation of Effects 3

This summary is based on the report:

Holmberg, I. Effekter av ny läroplan i matematik (Lgr 69) på vissa aspekter av talbegreppets utveckling, kunskaper i matematik och attityder mot ämnet. /Effects of a new curriculum of mathematics (Lgr 69) on certain mathematical concepts, and on certain achievement and attitude variables./ *Pedagogisk-psykologiska problem* (Malmö, Sweden: School of Education), No. 236, 1974.

2.1 Background and aims

The new curriculum, Lgr 69, was expected to lead to improvements in mathematics teaching. A number of factors influenced the form of the teaching:

Rapid social developments have made demands on the mathematics course. We have got a number of technical aids; computers, calculating machines, pocket calculators and slide rules. If we are to be able to use these aids effectively, new skills are needed. The students must be trained to handle the machines and check the results.

Statistics occupy a central position in our modern society. The newspapers daily supply us with diagrams and tables, the interpretation of which requires training, and statistics has become an expanded part of the course, with great practical value.

The demands made on knowledge of geometry have changed. The goal for teaching in geometry in the comprehensive school will be that the students should acquire basic geometric concepts, be able to handle various kinds of drawing equipment, be able to calculate circumference, area and volume, and be acquainted with symmetry, similarity and scale; knowledge which is useful in studying maps and plans.

The Swedish community's need for mathematics knowledge has changed. Consumer needs have been studied by Urban Dahllöf (1960, 1963). According to Dahllöf, upper secondary schools demand a high standard in the four rules

of arithmetic with integers, decimals and fractions, and in various kinds of application within geometry, algebra, commercial arithmetic etc. The need for a good knowledge of mathematics in occupations entered after the comprehensive school and in leisure time was considered to be slight—only the ability to count with integers was required.

Sven G. Jansson (1964) has arrived at rather similar results. He reports that arithmetic skills with whole numbers and the ability to make estimates largely covers the mathematics needed within the trades based on comprehensive schooling.

International research following new lines within e.g. mathematics methodology has contributed to the new curriculum. Some of the biggest projects have been discussed in the introduction to this thesis.

During the 1960's a worldwide investigation was carried out under the auspices of UNESCO. The mathematical achievements of school students in various countries were compared. Husén showed that Swedish 13-year-olds were worst in mathematics out of all the countries participating (Husén, 1967). This investigation has been criticized by Swedish mathematics teachers. According to a questionnaire answered by teachers, the mathematics tests on which the results were based were not considered adequate to Swedish conditions. Despite the criticism of the form of the tests, we can assume that the results of Husén's investigation intensified the demands for a modernization of the mathematics courses in Sweden, thus influencing the discussions of goals and objectives.

The discussion of goals and objectives (Nordic School Mathematics, 1967) has inspired the wording of the new directives. The goals of mathematics teaching were expanded to include the working habits of the student and his "maturity." The main emphasis in the goal for mathematics shifted from being instrumental (to prepare the students for tests by means of various assignments) to the intrinsic value of the subject. The goal should be to give the students a good ability in understanding numbers and problem-solving models, which can be transferred for use in many different branches of the subject.

The construction and evaluation of self-instructing and programmed material in mathematics has increased the desire for systematic goal-descriptions. The students' way of working with material-directed learning has also been analyzed and debated. This development is reflected in the Swedish discussions on goals and objectives.

The traditional way of working in mathematics teaching has been for the teacher to explain new items with the whole class and then for the students to work individually with a series of practice examples.

Group work has hitherto played a very small part in mathematics teaching. An experiment with small-group teaching combined with machine teaching

has shown satisfactory results (Holmberg, 1966). There is a great need for rational methods of cooperation in the subject to be tested and tried out. Individualized mathematics teaching has been tested on a large scale in the IMU project (Larsson, 1973 a and b). In all probability the goal-description of IMU has influenced the design for mathematics teaching given in Lgr 69 (cf. The Swedish Board of Education's handbooks: Basic skills in mathematics, 1973, p. 49). Taxonomic categories have not been used systematically, however, in the goal-descriptions on which study material adapted for Lgr 69 has been based.

The Nordic Committee for the Modernization of Mathematics Teaching conducted large-scale experiments with new study material during the first half of the '60's. Innovations such as basic operations in sets and the presentation of number systems other than the ten-system attracted widespread attention. Most important, however, was the introduction of new areas such as the theory of functions, vectors, and the theory of probability, together with the attempt at uniform presentation. *The goal-descriptions* in the two curricula are stated in general terms and are relatively similar. They both emphasize knowledge and counting skills.

Lgr 62, p. 164, Goals: The teaching in mathematics should develop the students' ability to handle quantitative concepts. The teaching should provide knowledge and skill in elementary arithmetic and algebra, together with familiarity with the elementary concepts and methods of geometry. On the basis of a clear understanding, the students should acquire confidence in solving various kinds of mathematical problems, primarily practical, by means of mental arithmetic and meaningful written exercises. The teaching of geometry should use the students' observations of the form of figures and bodies to train their ability in the perception of space and to develop their geometrical imagination. The students should gradually be made familiar with mathematical terms and forms of expression in general usage. The content of the mathematics teaching should open up to them with a wider perspective, on natural phenomena and society.

Lgr 69, p. 137, Goals: The teaching in mathematics is to work from the students' own experiences and ideas and to be based on understanding. It should gradually provide familiarity with some essential concepts and procedures in arithmetic, geometry, algebra and descriptive statistics, plus acquaintance with the concepts of function and probability. Further, the teaching is to train skill in numerical counting, including the use of technical aids, and to demonstrate how mathematics is used in different contexts.

The content of the mathematics courses in the two curricula is not radically different. No item in Lgr 62 has disappeared in Lgr 69. Magne, who was a member of the expert panel on the curriculum, describes the difference between Lgr 62 and Lgr 69 in the following way (Magne, 1973): "The main difference between the old and the new syllabus can be said to consist of a shift in emphasis on the following three points:

1. A desire to modernize mathematics, in order to adjust the school's mathematics teaching to the future needs of society.
2. Focus placed on the learning of concepts, in order to give the student "understanding" of basic concepts, of relations between them, and of the mathematical system of concepts.
3. Emphasis on differentiation of material and methods for different students, so that on the one hand clever students are given the opportunity of developing their ability, while the low-achieving students are protected from excessive demands."

In order to map the differences between the content of the two curricula, I have examined the content of Lgr 62 from the perspective of a goal-description that reflects the content of the mathematics syllabus in Lgr 69 (Holmberg, 1974, appendix 13). On the basis of this comparison the following innovations have been noted in Lgr 69:

Grade 3

1. Introduction of the symbols $>$ $<$ $\neq$
2. The number line used as a methodological aid
3. The use of the term algorithm: subtraction algorithm, multiplication algorithm
4. Theory of sets
5. Statistics

Grade 6

1. The number line used as a methodological aid
2. Expression of powers
3. Theory of sets
4. Theory of functions

Grade 9

1. Increased use of technical equipment such as sliderule, calculating machine and electronic computer

2. Theory of sets
3. Theory of functions; expanded course
4. Larger course in statistics—theory of probability
5. ADB, orientation

The teaching in accordance with the two syllabuses, using new study materials has, however, taken different forms. Mathematics teaching according to Lgr 62 has relied almost exclusively on class teaching and by tradition it has been accepted that the students should be able to use the four rules of arithmetic with some assurance. This assurance was acquired by practising many similar problems. "In the 1950's an elementary school student could well practise on 5,000 division problems between grades 3—8" (Magne, 1973 b, p. 3). Despite this intense training, skills in mathematics appear to have diminished successively (Vanås, 1953; Västgöta studien, 1972). The new syllabus was expected to make it possible to satisfy the demands that were being made for a better standard in mathematics. To get Lgr 69 off to good start, the supplement Lgr 69: II Ma was written, giving tips and guidance on the new syllabus. The comments are as a rule more detailed and concrete on the items that have not been included in mathematics in the comprehensive school previously. The new methodology that is recommended is based on the principle that all practice sections should be preceded by understanding. An attempt is made to make abstract ideas more concrete. In this laboratory material is used and the concept of sets is introduced. Detailed teacher's handbooks are produced in conjunction with all new study material and this is expected to produce more uniform teaching.

In-service training of teachers was initiated by the curriculum group in mathematics. It was hoped that this course would create the prerequisites for modern teaching of mathematics. The course (cf. Håstad, 1967 a and b) was given the name, the DELTA project. It was a project in which the Swedish Board of Education, the Swedish Radio Corporation and Hermods' correspondence school and publishing firm cooperated. The DELTA project involved about 45,000 teachers at all levels of the comprehensive school.

Courses for parents in the new mathematics were arranged. The intention here was to inform the parents of the purpose of the new mathematics and its content, in particular the theory of sets.

Despite these efforts to spread information about the new mathematics material, it proved impossible to introduce the new curriculum without arousing considerable criticism. After a year or two with the new syllabus, critical comments were heard concerning the new mathematics. These comments originated, all too often, from a lack of information about material and methods. It was felt purely subjectively that the students had a worse command of mathematics.

In a memorandum of 1970-11-21 the Swedish Board of Education asked for evaluation of the results of the teaching following Lgr 69 in certain subjects, including mathematics.

Parents who have not been familiar with the terminology of the new mathematics have asked many questions concerning the purpose and effect of the new system: Is this rearrangement of both material and teaching methods an adequate adaptation of mathematics to the demands made by the computer age on coordination of mathematics and logic? Do all categories of students, even those who are weak in mathematics, comprehend the reality that is abstracted in the terminology of sets? Can the students at such an early stage understand and experience the relation between the concrete operations they carry out and the terminology of the theory of sets, or is this two separate worlds for them? In the latter case, does this give rise to negative attitudes towards mathematics?

The main purpose of this project has been formulated in the context of the above discussion and is intended to study some effects of the introduction of a new curriculum in mathematics teaching.

2.2 Problems

The following problems have been extracted:

- A. Are there differences between student groups within different grades that have been taught according to Lgr 62 and Lgr 69 with regard to
- precision in the conception of number
 - knowledge of mathematics
 - attitudes towards mathematics as a subject?
- B. What influence does ability have on the above variables?

2.3 Methods

2.3.1 *Experimental design and methods of analysis*

Two groups of students, who have been taught according to Lgr 62 and Lgr 69 respectively have been *tested by means of test batteries*, which measure

1. precision in number conception (TAL 3, TAL 6 and TAL 9; presented in Holmberg, 1974, app. 9, 10 and 11).
2. knowledge in mathematics (Test papers MA 3, MA 6, MA9A and MA9S; presented in Holmberg, 1974, app. 5, 6, 7 and 8).

3. attitudes towards the subject of mathematics (Attitude MA; presented in Holmberg, 1974, app. 12).
4. ability (Test Cattell Culture Fair 2 A has been used for grades 3 and 6; Cattell Culture Fair 3 A has been used for grade 9; Cattell, R. B. & Cattell, A. K. S., 1965).

The tests were administered in week 16 of 1972 for student group A and in week 15 1973 for student group B. The class teachers in grades 3 and 6 and the mathematics teachers in grade 9 gave all the tests. They had been instructed by the project leader and had been given detailed directives to follow.

In the analysis of the material, the students have been divided up into three levels of ability, high, middle and low groups, chosen on the basis of the results from the ability test Cattell Culture Fair 2 A and 3 A. Statistical analysis has been carried out by means of 2×3 factorial analysis (ANO 2 P, Sternerup, Lund Computing Center). Factor A is fixed and has two levels; Lgr 62 and Lgr 69. Factor B is also fixed and has three levels, corresponding to the different levels of ability. In order to determine the validity of the instruments, correlations have been calculated: product correlations and canonical correlation analyses (Dixon, W. J., 1968). The reliability has been calculated with Cronbach's α coefficient (Lord, F. M. & Nowick, M. R., 1968) and Pearson correlations.

Analyses and processing in this investigation have been carried out on the computer UNIVAC 1108, at the Lund Computing Center.

In the analyses of variance ω^2 has been included in order to show how great a proportion of the total variance a particular difference is responsible for (cf. Hays, 1963, p. 406—407, 425). In these analyses of variance $\omega^2 \geq 0.05$ has been made the criterion for whether differences between the groups are significant.

2.3.2 *Subjects*

Group A: about 200 students from grades 3, 6, 9 in the school year 1971/72 represented students taught in accordance with the 1962 curriculum.

Group B: about 200 students from grades 3, 6, 9 in the school year 1972/73 represented students taught in accordance with the 1969 curriculum.

In the autumn of 1970 grades 1, 4 and 7 of the comprehensive school started working with Lgr 69. The syllabuses were introduced successively and in 1973 everybody in the comprehensive school was working with "the new mathematics." For this investigation that means that the students in grade 3 have had the same curriculum all the time they have been at school, while grades 6 and

9 in experiment group B changed curriculum after 3 and 6 years respectively with the old one. The students in these classes have been subjected to the strain that always accompanies changes in the school's outlook and methods of work.

For the purposes of generalization it would have been desirable to have a random sample of students, but for economic and practical reasons this proved impossible. An occasional sample has been chosen on the basis of the following logical arguments. The variables that are important for my population are home town, social background and intelligence and these must be equally represented in both experiment group A and B.

The students have been taken from different school districts, which represent both town and country. Each experiment group consists of 50 % from town and 50 % from the country. All students in grades 3, 6 and 9 at the chosen experiment schools have been included. Group A consists of grades 3, 6 and 9 in the school year 1971/72 and Group B of grades 3, 6 and 9 at the same schools in the school year 1972/73. Thus the same schools have been used as experiment schools in both 1971/72 and 1972/73 and this factor is in our opinion a guarantee that we have got socially comparable groups. The children in Groups A and B belong in this way to the same intake area. The two experiment groups are intellectually comparable, tested by means of Cattell Culture Fair 2 A and 3 A respectively ($p > 0.05$, $\omega^2 = 0.01$).

When the attendance at our schools is compared to that at other schools (The SIA investigation "Absence from School"; Nilsson, 1974; SIA concerns the internal work of the school, and "School Statistics"), our experiment schools prove to reflect the normal attendance figures in the comprehensive school. (Percentage of absentees on the test occasions: grade 3 = 2 % both school years; grade 6 = 3 % and 4 % respectively; grade 9S = 5 % and 6 % respectively; grade 9A = 8 % and 10 % respectively).

These student groups, which were tested in our 1973 study, were the first groups that the teachers had had with the new curriculum. Their unfamiliarity with the new order can naturally have influenced the results of the investigation negatively.

2.3.3 *Teachers*

In grades 3 and 6 qualified class teachers have been responsible for the teaching in all the experiment classes. All the teachers have had at least three years' experience at the school level in question. In the upper level classes subject teachers with academic degrees have done the teaching in mathematics. The teachers using Lgr 69 have had in-service training: the DELTA project mentioned earlier.

2.3.4 *Study materials and activities (teaching methods)*

The two experiment groups have used different study materials. The Lgr 62 course had a textbook in mathematics with few instructions and an abundance of practice examples for each step of the course. The students had to copy the items and calculate them in separate exercise books. Lgr 69 inspired several textbooks, teachers' handbooks and experiment material. The students are allowed to write directly in the study material and only copy the exercises at certain points.

The teachers consider that the study materials for Lgr 69 provide a high degree of concreteness, since it includes e.g. multi-base blocks and a geometry box. The quality of the new study materials naturally has considerable effect on the results of the investigation.

Both Group A and Group B have had the same amount of homework.

The two experiment groups have had the same organizational model: collective teaching (class teaching). No experiments have been made with large class teaching or individualization according to the IMU model at any level. The teachers point out, however, that the new study materials have controlled their methods more strictly: more individualization has been possible because of the design of the study materials. The forms for planning the work was the same for both groups. The teachers in grades 3 and 6 planned individually together with the students, while the grade 9 teachers planned in common for the whole school year.

2.3.5 *Instruments*

In order to illuminate the main problems in the investigation we need both achievement tests: number comprehension tests and knowledge tests in mathematics for grades 3, 6 and 9, and attitude tests.

2.3.5.1 Number conception tests

Number conception is part of mathematical ability. It is grounded when the students handle numbers—compare them, combine them etc. It is not sufficient to present the numbers to the student and then expect him to have a well-grounded conception of numbers. Good number conception implies the ability

- to identify: recognize number
- to generalize: collect in classes; e.g. one, two and three figure numbers, positive and negative numbers, integer and decimal numbers, rectangles, squares etc.
- to think in terms of relations $6 < 7 < 8$; $22 = (7 \times 3 + 1)$.

Piaget claims that children acquire number conception partly through a growing ability to reason logically and partly by manipulating numbers and objects.

Work done by Piaget, (1969), Wohlwill, (1960, 1962), Lovell, (1971), et al. has formed the basis of the project leader's construction of the number conception tests TAL 3, TAL 6 and TAL 9. Reliability is stated with Cornbach's α coefficients (Lord & Nowick, 1968):

Grade 3: $\alpha = 0.85$; grade 6: $\alpha = 0.81$; grade 9: $\alpha = 0.86$.

Since number conception is part of mathematical ability, the relation number conception/knowledge in mathematics can be said to illustrate the validity of the number conception test. The validity of the number conception tests is shown in canonical correlation analyses. A canonical analysis can be regarded as a linear regression model, in which the dependence of one group of variables on the other is studied (Mårdberg, 1969, p. 132). Canonical correlations:

Grade 3		Grade 6		Grade 9	
Lgr 62	Lgr 69	Lgr 62	Lgr 69	Lgr 62	Lgr 69
0.86	0.86	0.85	0.84	0.82	0.91

It seemed important to validate these new number comprehension tests with a test that had not been produced internally in the project. Therefore a further measure of the validity of the number comprehension tests is given in the form of the relation between number conception and mathematical understanding, as it is measured in parts of the standard tests for grades 3 and 6.

Pearson correlations:

Grade 3		Grade 6	
Lgr 62	Lgr 69	Lgr 62	Lgr 69
$r = 0.62$	$r = 0.72$	$r = 0.72$	$r = 0.70$

Since no standard tests are given in grade 9, the TAL 9 test cannot be validated in the same way.

2.3.5.2 Knowledge test in mathematics

The test papers called MA 3, MA 6, MA 9A (general course) and MA 9S (special course) have been used to compare the knowledge in mathematics of students taught in accordance with Lgr 62 and Lgr 69. The tests are presented

in Holmberg, 1974, p. 20 and app. 5—8. The same test papers have been given to students taught in accordance with Lgr 62 and Lgr 69 respectively. Only items that occur in both curricula have been included in the tests, which consist of a cross between Lgr 62 and Lgr 69. The goal-descriptions for both curricula are vague. When constructing tests, one nevertheless relates the test to the goal-description and tries to find tasks that can be criteria for that the evaluation is adequate with regard to the goal. The effects of the innovations within Lgr 69 have only been tested in so far as they have had transfer effects on other sections of the mathematics test.

Each year Sweden can measure the effect of mathematics teaching under the current curriculum by means of the standard tests. In this investigation it was important to construct instruments that related well to the standard tests of both 1972 (referring to Lgr 62) and 1973 (referring to Lgr 69).

Therefore the results of the standard tests have been used as a criterion of the validity of the mathematics tests for grade 3 and grade 6. The validity coefficients vary between $r=0.82$ and $r=0.40$, Md: $r=0.69$. The tests for grade 9 have been validated against the students' grades in mathematics, $r=0.65$. The validity coefficients can be considered satisfactory for both curricula.

The reliability has been calculated with Cronbach's α coefficient (Lord & Nowick, 1968).

Reliability coefficients:

MA 3 $\alpha=0.88$

MA 6 $\alpha=0.86$

MA 9A $\alpha=0.89$

MA 9S $\alpha=0.86$

2.3.5.3 Test of attitudes towards mathematics: Attitude MA

Attitude MA is an instrument constructed by the project leader for measuring the attitude of school students towards mathematics. It is a Likert-type scale. The students must consider 14 statements and in each separate case mark one out of five alternative answers. The validity has been determined indirectly. Here the differences in attitude between groups with extreme variations in grades have been tested. The high group has significantly higher attitude points than the low group ($p<0.01$, $\omega^2=0.07$). The reliability has been calculated by means of the test-retest method (Pearson correlation), $r=0.62$.

2.4 Results and discussion

2.4.1. *Analysis of number conception*

Grade 3

No significant differences can be demonstrated in this analysis between the effect of the two curricula on the conception of numbers. The talent variable provides a significant contribution to variance ($p < 0.01$, $\omega^2 = 0.20$). The high group is best at number comprehension and the low group worst. This result is obtained in all grades and appears natural. No significant interplay can be noted.

Grade 6

Students taught according to Lgr 69 have significantly better results in the number conception test than those who studied the old mathematics ($p < 0.01$, $\omega^2 = 0.07$). This applies on all levels of ability. The talent variable has a significant effect ($p < 0.01$, $\omega^2 = 0.26$). No significant interplay occurs.

Grade 9

No differences can be demonstrated between the ability of the curricula to influence number conception in grade 9. The effect of the different levels of talent is significant ($p < 0.01$, $\omega^2 = 0.36$). A significant interplay is noted ($p < 0.05$, $\omega^2 = 0.01$). The interaction implies that while the high and middle groups show the best results following Lgr 69, the low group is best when following Lgr 62. However ω^2 is only 0.01, and therefore the effect is of no real importance.

2.4.2 *Analysis of knowledge in mathematics*

Grade 3

The students taught according to Lgr 69 with the so-called new math have significantly better results on the MA 3 test paper than those taught according to Lgr 62 ($p < 0.01$, $\omega^2 = 0.02$). This applies to all levels of ability. However ω^2 is too low to permit generalizations on the basis of the effects. In mathematics too the effect of the different levels of ability is significant in all grades ($p < 0.01$, $\omega^2 = 0.11$).

Grade 6

On all levels of ability students taught according to the so-called old mathematics have significantly better results than those taught according to Lgr 69 ($p < 0.01$, $\omega^2 = 0.05$). The results for grade 6 are the opposite of those for grade 3.

Grade 9, general course

Students taking the general course who have been taught in accordance with Lgr 62 manage the mathematics test better than those who have been taught according to Lgr 69 ($p < 0.01$, $\omega^2 = 0.13$). This applies to all levels of ability.

Grade 9, special course

Taken as a whole no differences in mathematics achievements can be shown between the two curriculum groups.

2.4.3 *Analysis of attitudes towards mathematics*

Grade 3

No differences in attitude can be demonstrated between students taught by Lgr 62 and Lgr 69 respectively. The difference in attitudes between the levels of ability is not significant at this age-level.

Grade 6

The attitudes towards mathematics appear to be most positive in those students taught according to Lgr 62 ($p < 0.01$, $\omega^2 = 0.03$). Significant variation can be traced between the levels of ability ($p < 0.01$, $\omega^2 = 0.04$). In no case is the variation contribution large (cf. ω^2).

All students in grade 9

No significant difference in attitudes towards mathematics exists between the two curriculum groups. On the other hand a significant variation can be noted between levels of ability ($p < 0.01$, $\omega^2 = 0.28$). The low group has the most negative attitudes towards mathematics and the high group the most positive.

Grade 9, general course

No significant variation can be demonstrated, nor interplay with other variables.

Grade 9, special course

The curriculum has no effect on attitudes towards mathematics. Talent on the other hand has such an effect ($p < 0.01$, $\omega^2 = 0.06$). The low group in 9S has a markedly more negative attitude towards mathematics taught according to Lgr 69.

All students

All the students participating in the investigation have answered the same attitude test and it has been possible to analyze all the results.

Totally there are no differences in attitude towards mathematics depending on which curriculum has been followed in the teaching.

The grade in question influences the attitude significantly. Grade 3 is most positive and grade 9 most negative towards mathematics as a subject ($p < 0.01$, $\omega^2 = 0.21$). A significant interplay ($p < 0.01$) can be shown, but ω^2 is low (0.01). Grade 3 is most positive towards the new mathematics, while grades 6 and 9 are most positive towards the old one.

2.5 Summary

1. The analysis of the *ability in number conception* shows that students in grade 6 who have been taught according to Lgr 69 achieve significantly better results than those who studied the old mathematics. For students in grades 3 and 9 this analysis shows no significant differences between the effect of the two curricula on number conception. The talent variable gives a significant variance contribution in all grades. The high group achieves the best results in the number conception tests and the low group the worst results. A significant interplay is noted for grade 9. The interaction implies that while the high and middle groups get the best results when following Lgr 69, the low group does best following Lgr 62. This result should be interpreted with caution, however, since ω^2 is low ($\omega^2 = 0.01$).

2. The analysis of *knowledge in mathematics according to our investigation* shows that students in grade 3 who have been taught according to Lgr 69—the so-called new math—succeed as well as those in grade 3 who have followed Lgr 62—the so-called old math.

In grade 6 students who have been taught by the old mathematics achieve better results.

In grade 9 the students take either a general or a special course in mathematics and for this reason two separate tests and analyses have been made. Students taking the general course in mathematics succeed best with the old mathematics. For the students taking the special course no differences

can be demonstrated in mathematics achievements as a result of which curriculum has been followed.

In all the grades the high group has the best knowledge of mathematics and the low group the worst. No significant interplay occurs.

3. *Attitudes towards mathematics* have been tested with the same instrument for all grades. Totally there are no differences in attitudes towards mathematics between the two curriculum groups. The students in grade 3 are more positive towards mathematics than the students in grade 6, who in their turn are more positive to mathematics than those in grade 9. In grade 3 there is no difference in the attitudes of the high and low groups, while in grades 6 and 9 the high group has more positive attitudes than the middle and low groups.

2.6 Conclusions

Ultimately it is the study materials and the teaching—an operational expression of the curricula—that have been evaluated in this investigation. It should be remembered that it is only one section of the two curricula that has been tested. No direct test has been made of the effect of the innovations to be found in Lgr 69. Naturally students taught according to Lgr 69 must know more about e.g. sets and the theory of function than students who have followed Lgr 62 and consequently never been taught these aspects of mathematics.

In our increasing technical society mathematics teaching must provide the students with both a rapid conception of numbers and good mathematical skill. When the new course was designed, more time and space was allotted to concept formation and systematics, which had a positive effect.

Lgr 69 has not been equally successful in imparting knowledge in grade 6 and grade 9, general course. The arithmetic skills of these students should be improved. Mathematics is largely a question of skill. A mechanical error in a calculation can lead to completely misleading results, without this being immediately apparent from the context. The teaching must therefore give the students the opportunity of practising their skill in arithmetic, so that when they leave school they can make calculations with some degree of self-confidence. Thus the social competence of the students must be developed further in mathematics.

3. Final discussion

In the ten years covered by these investigations of effect, the goals, content and methods of Swedish mathematics teaching have undergone important changes. International conferences and research projects have stimulated both the Swedish debate on goals and mathematics research.

At the beginning of the '60's teaching machines and programmed teaching PI were very much the topic of the day and it was only natural that projects were started to investigate their potential. In our investigations—an intensive observation study lasting six terms—interest was focussed on the reactions of the student when he was confronted by a teaching machine and program. We wished to see how the student learned mathematics when the subject was structured according to Skinner's principles of learning. We could note the importance of

- the student being allowed to work at his own pace
- the material being presented in small steps
- the material being well-sequenced
- the student working actively with the material
- the student receiving reinforcement—being given the right answer immediately.

The combination of all these principles into a learning strategy was then considered an outstanding educational finesse. The machine that presented the program was thought of as a substitute for the teacher and my results showed the machine to be an acceptable substitute in providing knowledge. What the students in the machine class missed was the social interplay with the teacher: a teacher function that the machine could not provide. Another negative effect of PI was the monotonous work routine: the students considered that they felt more tired with PI than with conventional teaching. In return many felt that they learned more. We have not been able to verify this opinion by knowledge tests. Both methods produced equally good achievements in mathematics.

Alternating methods appears to favor the learning climate. Students who have had machine teaching varied with small-group teaching caused less disturbance than students who had conventional teaching throughout.

The teacher in the machine class thought that a maximal use of teaching machines gave him an uninteresting role as a distributor of material. The modified form of the experiment made the teacher's work more stimulating, which was appreciated. In the debate of effectiveness that developed in the wake of teaching machines and programmed material, the value of the co-operation between teacher and students was overlooked. Bruner realized this and wrote: "In organizing our educational system we have been curiously blind to the extent to which knowledge is based on mutual dependence. We have 'teachers' and 'students', 'experts' and 'laymen'. But the fact that knowledge implies fellowship has somehow been overlooked" (Bruner, 1966).

We noted in our experiments that the students in the machine class increased the effectivity of their work—they could read instructions and carry out laboratory experiments more independently than the parallel class. In relation to the goal-description given in Lgr 62 this was a very successful result. "In order that the work should proceed smoothly and the students acquire good working habits, they should learn from the very first grades to work according to given instructions. The demands made on their independence and responsibility increase in grades 4—6." (Lgr 62, p. 61.)

In 1965 we terminated our project with investigations into programmed material in mathematics. The IMU project became the new mathematics project at the Department.

A large number of research institutes have continued to work on the possible uses of programming techniques. The focus of interest has shifted from the student who learns to the programmer who teaches. Skinner's principles have been tested and these research results have formed the basis for new programs and the production of other study materials. Some more recent research results concerning Skinner's principles are commented on below.

Working at one's own pace without teacher guidance produces a positive reaction at first, but after a term or two of working only with a machine the students express a desire for more teacher contact (Holmberg, 1965). "... research has indicated that the teacher and the program working together provide a more efficient teaching technique than either working in isolation" (Hartley, 1974, p. 280). Administrative difficulties arose when the students worked at their own pace. The spread within a group of students could be considerable—different students took different lengths of time to complete a program. Nowadays one tries to alternate individual work with other methods.

The *small-step technique* proved to be not always the best in all learning situations. Holland (1965) and Leith (1966) present some points to be observed when choosing the size of the steps:

- university students find small steps irritating
- small steps are most suitable for small children
- small steps are suitable for students who are afraid of a subject
- small steps are good in subjects that are logically constructed, e.g. statistics.

At the end of the '60's programs were made with varying step sizes (cf. Markle, 1969).

The research into *sequencing* was preoccupied at the beginning of the 1960's with the choice of flow model. The Skinner model (a simple linear model in which all students follow the same sequence through the material) was compared to the Crowder model (a branching model with different sequence through the material depending on the answers given by the student). Cram (1961) shows that if the subject material has a firm structure, the Skinner model is preferable, while branching programs are better suited for subject material with comparisons and evaluations. Later the research on sequencing has concentrated on assessing Gilbert's chain model (Gilbert, 1962; Mechner, 1967), Mager's student-paced sequence model (Mager, 1961; Mager & Clark, 1963) and sequencing according to Gagné's learning hierarchy (Gagné, 1970; Gagné & Briggs, 1974). In Gilbert's model the components are "chained together" in a program by the right answer to one task being incorporated into the next task unit as one or more words in different type, instead of being presented separately. The superiority of this model has not been proved (Wilcox, 1974). On the other hand a number of studies have shown that Mager's student-paced sequencing is superior to teacher-pacing (Horn, 1964; Briggs, 1968; Newkirk, 1973). Gagné maintains that some subjects have a hierarchical structure and must be sequenced accordingly. The results of the research on Gagné's sequence model are not unequivocal, but the model is still being investigated (White, 1974).

Active responding proved to produce better results with younger children (Holland, 1965; Anderson, 1967).

Immediate knowledge of results has proved to be most valuable for students of low ability and for students working with difficult programs that lead to a high frequency of errors (Holland, 1965; Annett, 1969). Tait et al., 1973, show that students who correct their mistakes and get reinforcement gradually achieve better results than the students who receive global reinforcement.

In *evaluating* programs interest was formerly focused primarily on the terminal results, but today every phase of the learning process is studied (Cooley, 1973).

Programmed material is today produced mostly for special courses in order to save on teacher resources.

Skinner's principles have been shown to have their limitations but to function satisfactorily in many contexts.

The concept of programmed instruction has widened since its early beginnings. Hartley (1974) has indicated the changing nature of programmed instruction in the following way:

1954	1961	1968	1974
	behavioural objectives task analysis	task analysis behavioural objectives subject analysis	systematic analyses and constraints levels of objectives techniques of structuring material
small steps overt responses	small steps logical sequencing	task structures and related teaching strategies	task structures and related teaching methods, strategies
reinforcement	active responding	presentation as a communication problem	and tactics presentation as a communication problem; problems of learning from text, diagrams, etc.; appropriate instructional decisions
self-pacing	immediate feedback self-pacing		individual differences and methods of learning
revision	evaluation (i) revision (ii) comparison studies	evaluation: revision	evaluation of (i) objectives and methods (ii) techniques: retrospective and concurrent decision making evaluation as illumination
			installation and implementation inservice teacher training

(each cell indicates areas of research)

Notes:

- (i) How the issues have been seen to get more complex as time has passed—so that research has spread into new areas.
- (ii) How research on programmed instruction and instruction per se have now come almost to overlap.

The essence of modern programmed instruction appears to lie in four complexes and closely-interrelated steps: These are:

- (i) knowing where you are going (specifying objectives),
- (ii) trying to get there by the best means possible (interrelating task analyses and teaching strategies),
- (iii) assessing whether you have got there successfully (evaluation), and
- (iv) using the results of this assessment to improve the methods used in step (ii), (revision).

This wider concept of programmed instruction permits the use of a variety of methods and offers a new approach to teaching. When this approach is employed on a large scale, computer-assistance is necessary, either for the teaching itself or for its organization.

The interest in teaching machines cooled at the beginning of the 1970's. Mathematics programs with computer-assisted instruction has not spread as much as was hoped at the end of the 1960's. Perhaps there will be a change on this point in the future. The teacher-shortage in Germany is being compensated just now by means of LIDIA (Lernen In DIAlog), a computer-assisted mathematics program. In this experiment, which is underway in Bayern, teacher-led computer instruction is being used to provide the students with human contact. Thus the educational technological system of the 1960's has developed into a "pedagogical praxis," which means that the teacher leads the students into a relationship with the material that involves not only learning, but also conversation, reflection and interpretation.

The growth of mathematics and a new view of the uses of mathematics as a technical instrument has created a new context for the content of school mathematics. The objective of mathematics teaching has been changed from an instrumental use of mathematics towards teaching the subject for its own sake. In Sweden the new syllabus in mathematics was introduced in Lgr 69. The third part of this thesis describes the effect of the introduction of this syllabus in mathematics teaching with regard to

the development of number conception,
knowledge in mathematics,
attitudes towards mathematics.

The effects of teaching mathematics according to Lgr 69 were neutral for grade 3 in all the investigation variables, positive for grade 6 in number conception but negative in mathematics knowledge and attitudes, neutral for grade 9 general course in number conception but negative in mathematics knowledge, neutral for grade 9 special course, all being taken in relation to the effects of teaching according to Lgr 62.

In *grade 3* the students had no knowledge of earlier mathematics methods and enjoyed following Lgr 69. The teachers on the junior level had been better prepared for Lgr 69 than many other categories of teachers. For a number of years they had been involved in an intense debate on methods and material (Bergfeldt, 1973). The study materials for the junior level had previously been thoroughly restructured.

In *grade 6* knowledge of mathematics decreased with Lgr 69. This can perhaps be because the grade 6 students had been forced to abandon well-known counting habits and had become uncertain when carrying out mathematical operations. The uncertainty resulted in negative attitudes.

The greater demand for understanding which the new method requires has, however, led to an improvement in number conception.

In *grade 9* the students taking the special course did well in the new mathematics, but the general course achieved worse results than with the old methods. Number comprehension was better with Lgr 69 for both high and middle groups, irrespective of whether they had the special or general course, while the low group demonstrated a more assured number comprehension with Lgr 62. Difficulty in learning the new terminology and new course items can have had the greatest effect on the results of the low group.

In 1971 the LUG group (Swedish: *LäroplansUtveckling inom Grundskolan*; Curricular development in the comprehensive school) in mathematics decided to investigate the mathematical skills of comprehensive school students. This investigation reports improved mathematical skills for the junior level and a deterioration for the middle and upper levels (Eliasson, 1973). The results of our investigation support Eliasson's results for grade 3 and to some extent for grade 6 too.

We can show one positive effect of Lgr 69 in grade 6: the students have better number conception. Our investigation gives a more differentiated picture of mathematics proficiency in grade 9 than Eliasson's, since in the analyses of variance we can extract the influence of the level of achievement. In the 1974 LUG investigation a continued improvement in grade 3 was reported, somewhat better results for grades 6 and 9A and the same level as before for grade 9S.

When introducing new material and new working methods in the middle of an education process, a consequent drop in skills can appear in the initial phase. A similar development has been reported from New Hampshire (Austin & Prevost, 1972). When new mathematical material was introduced into the courses, there was a decline in mathematical skills but at the same time mathematical understanding increased to above the level that had been normal before the change of syllabus.

The effects of the new elements in Lgr 69 have not been tested directly. It is only the items that are to be found in both the curricula that have been tested. It is natural that students taught according to Lgr 69 know more about e.g. sets and the theory of function than students taught by Lgr 62, which never took up these items. Probably Lgr 62 has been favored in the comparison with Lgr 69, since it has taken time to introduce the latter. The effects of the innovations in Lgr 69 appear in our investigations only as transfer effects.

In our increasingly technical society mathematics teaching must provide the students with the ability to comprehend number quickly and with solid mathematical skills. Lgr 69 has succeeded in giving the student better number conception. The important task now is to continue research into ways of increasing the mathematical skills of the students.

The Nuffield project in England established that the new mathematics did not give the students the skill in calculating that they needed in other subjects, such as physics and chemistry. The teachers were of the opinion that the content of the mathematics courses had become so large that the training of skills was neglected. A new project known as SUM (Science Uses Maths) is now studying how mathematics can be used in other subjects and how different subjects can be integrated (Times Educ. Supp., 73.10.05).

Perhaps a joint approach would enable the student to proceed successfully with further studies and employ his mathematical knowledge efficiently.

The curriculum reform movement in school mathematics was aimed primarily at improving goals, content and teaching procedures for the subject. Nevertheless, the movement did also stimulate research into mathematics teaching. The rapidity with which the school mathematics curriculum has changed and the recognition that the curriculum will now require continual updating has inspired increased research on mathematics education in many countries. The main thrust has been in "applied research"—chiefly the development and testing on new methods and materials. Comparative studies of different instructional programs have become a common type of research in mathematics teaching. Most of these studies have been open to criticism on both conceptual and methodological grounds. Some authorities, including Romberg (1969) have argued that evaluation studies are not research at all. However, even the results of evaluation studies can be a driving force that can lead to changes in mathematics teaching. The criticism mentioned above has initiated a number of new trends in educational research, designed, conducted and reported in such a way as to secure improved education in the classroom.

This new strategy can be described as "developmental research." In such research the investigator explores various ways of treating a particular topic in the curriculum. The information obtained about the relative advantages and disadvantages is continually feed back to the teachers and the curriculum developers.

A popular approach in developing educational models has been deduction from learning models. Researchers in mathematics education have borrowed as much as they could from Piaget's developmental theory, Skinner's reinforcement theory, Gagné's theory of learning by discovery, Carroll's model of mastery learning etc.

This procedure has also been attacked. Psychologists such as Hilgard and Bugelski (cf. *New trends in mathematics teaching*, 1972, p. 133) have asserted that current learning theories have obtained most of their evidence from animal experiments and therefore have little to do with the teacher's chief resource: language.

One of the latest trends involves descriptive presentation of data. Descriptive

studies that follow the minutiae of classroom interplay have met with some dissatisfaction. There appears to be a new tendency in which an individual teacher's minute-by-minute decisions and strategies in the classroom are studied (cf. Lundgren, 1973).

Computers are being used increasingly to gather data on learning and teaching; computer-assisted research studies have been started.

A movement towards a unified mathematics curriculum from the first school year to university level has implications for future research activities. Recent research has shown that even young children can be taught rather advanced subjects. Why should such subjects be taught earlier and what are the effects of doing so? More studies will be required of the implications of sequencing material in new ways. A unified curriculum demands a complete rethinking of the school program, supported by research evidence.

Modern mathematics, with its invasion of all the sciences, must include examples of the newer applications in the school curriculum. As the content changes, there is an accompanying change in teaching methods and materials for instruction. The search for new ways of improving mathematics teaching and learning must continue.

References

- Anderson, R. C. Educational psychology. I: *Annual Review of Psychology*, 1967, 18, 129—164.
- Annett, J. *Feedback and Human Behaviour*. Harmondsworth: Penguin, 1969.
- Austin, G. R. & Prevost, F. The result of a longitudinal study in mathematics in New Hampshire. *Math. Teach.*, 1972, 65, 746—747.
- Barson, A., Cochran, B. & Davies, R. Childcreated mathematics. *Arithmetic Teacher*, 2, 1970.
- Basfärdigheter i matematik*. SÖ:s handledningar, Stockholm, 1973.
- Begle, E. G. (Ed.) *Mathematics education. 69th Yearbook NSSE, Part I*. Chicago: University of Chicago Press, 1970.
- Bergfeldt, S. E. *Ökar laborativt material matematikfärdigheten på låg- och mellan-stadierna?* Linköping: Lärarhögskolan, 1972 (Stencil).
- Bjerstedt, A. *Självinstruerande studiematerial*. Lund: Uniskol, 1963.
- Bjerstedt, A. *Educational technology*. Lund: Gleerup, 1972.
- Björk, L.-E. Effekter av BASIC-programmering i gymnasieskolans åk 1 på elevernas numeriska färdigheter och attityder mot matematik. *Pedagogisk-psykologiska problem*. Malmö: School of Education. No 246, 1974.
- Briggs, L. J. *Sequencing of Instruction in Relation to Hierarchies of Competence*. Pittsburgh: American Institutes for Research, 1968.
- Bruner, J. S. *På väg mot en undervisningsteori*. Lund: Gleerup, 1966.
- Bruner, J. S. *Undervisningsprocessen*. Lund: Gleerup, 1970.
- Cattell, R. B. & Cattell, A. K. S. *Culture Fair Intelligence Test Scale 2*. Champaign, Ill.: Institute for Personality and Ability Testing, 1965.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press, 1969.
- Cooley, W. W. *Evaluation in a research and development context*. Paper from the Learning Research and Development Center, University of Pittsburgh, 1973.
- Cornell, F. G., Lindvall, C. M. & Saupe, J. L. *An exploratory measurement of individualities of schools and classrooms*. Urbana: Bur. Educ. Res., Univer. Illinois, 1952.
- Coulson, J. E. & Silberman, H. F. Effects of three variables in a teaching machine. *J. educ. Psychol.*, 1960, 51, 135—143.
- Cram, D. *Explaining "teaching machines" and programming*. San Francisco: Fearon, 1961.
- Crowder, N. A. Intrinsic programming: Facts, fallacies and future. I: Filep, R. T. (Ed.) *Prospectives in programing*. N. Y. Macmillan, 1963, 84—115.
- Davies, I. K. & Hartley, J. (Eds.) *Contributions to an educational technology*. London Butterworths, 1972.

- Dahllöf, U. *Kursplaneundersökningar i matematik och modersmålet*. (SOU 1960:15) Stockholm, 1960.
- Dahllöf, U. *Kraven på gymnasiet*. (SOU 1963:22) Stockholm, 1963.
- Deterline, W. A. et al. Programmed instruction today. *Educational Technology*, 1970, 10 (7), 29—47.
- Diagnostiska Läs- och Skrivprov (DLS)*. 2:a uppl. Stockholm. Skand. Testförlaget, 1964.
- Dixon, W. J. (Ed.) *BMD Biomedical computer programs*. Berkeley: University of California Press, 1968.
- Edwards, A. L. *Experimental design in psychological research*. New York: Holt, Rinehart & Winston, 1960.
- Eigen, L. D. High-school student reactions to programed instruction. *Phi Delta Kappan*, 1963, 44, 282—285.
- Eliasson, L. LUG: *Undersökning av elevernas räknefärdighet inom grundskolan*. Falun: Länsskolnämnden i Kopparbergs län, 1973. (Stencil.)
- Eliasson, L. LUG: *Undersökning av elevernas räknefärdighet inom grundskolan*. Falun: Länsskolnämnden i Kopparbergs län, 1974. (Stencil.)
- Evans, J. L., Glaser, R. & Homme, L. E. A preliminary investigation of variation in the properties of verbal learning sequences of the "teaching machine" type. I: Lumsdaine & Glaser (Eds.), *Teaching machines and programed learning: A source book*, 1960, 446—451.
- F-provet: Form A*. Stockholm, Skand. Testförlaget, 1960.
- Feldhusen, J. F. Tapes for teaching machines. *Phi Delta Kappan*, 1963, 44, 265—267.
- Flinck, R. *Skolanpassning hos elever på enhetsskolans/grundskolans högstadium*. Rapport från pedagogiska institutionen, Lunds universitet, Lund 1965. (Stencil.)
- Flinck, R. *Lundasystemet i matematik. Konstruktionsarbete, systemets uppbyggnad, resultat av lärarenkäten I*. Lund, 1970.
- Frank, H. (Ed.) *Lehrmaschinen in kybernetischer und pädagogischer Sicht*. III. München: Verlag Oldenbourg, 1965.
- Gage, N. L. (Ed.) *Handbook of research on teaching*. Chicago: Rand McNally, 1964.
- Gagné, R. M. & Dick, W. Learning measures in a self-instructional program in solving equations. *Psychol. Rep.*, 1962, 10, 131—146.
- Gagné, R. M. *The Conditions of Learning* (2nd ed.). New York: Holt, Rinehart and Winston, 1970.
- Gagné, R. M. & Briggs, L. J. *Principles of Instructional Design*. New York: Holt, Rinehart and Winston, 1974.
- Gilbert, T. F. Mathematics—the technology of education. I: Merrill, M. D. (Ed.) *Instructional Design: Readings*. Englewood Cliffs: Prentice Hall, 1971.
- Glaser, R. & Taber, J. I. *Investigations of the characteristics of programed learning sequences*. Pittsburgh: Programed Learning Laboratory, Univer. of Pittsburgh, 1961.
- Gran, B., Gran, Bi., Rudvall, G. & Röhr, G. Några pågående forsknings- och utvecklingsarbeten i Malmöregionen med relevans för SIA:s förslag. *Pedagogisk-psykologiska problem*, Malmö: School of Education, No 259, 1974.
- Gropper, G. L. & Lumsdaine, A. A. An investigation of the role of selected variables in programed TV instruction. *Studies in Televised Instruction, Report No. 4*. Pittsburgh. Metropolitan Pittsburgh Educational Television Stations WQED—WQEX and American Institute for Research, 1961.
- Hall, T. *Matematikens utveckling*. Lund: Gleerups, 1970.
- Hartley, J. Programed instruction 1954—1974: A review. I: *Programed learning & educational technology*, 1974, 11 (6), London.
- Hays, W. L. *Statistics*. New York: Holt, 1963.

- Holland, J. G. Research on programming variables. I: Glaser, R. (Ed.) *Teaching Machines and Programed Learning*, II. Washington: N. E. A., 1965.
- Holmberg, I. Effekten av programmerad undervisning på vissa prestations-, attityd- och beteendeveriable, jämfört med konventionell undervisning. *Pedagogisk-psykologiska problem*. Malmö: School of Education, No. 18, 1965, (Rev. uppl., 1975).
- Holmberg, I. Maskinundervisning kombinerad med lärarledd smågruppsundervisning, jämfört med konventionell undervisning. (Effektundersökning i matematik på högstadiet). *Pedagogisk-psykologiska problem*. Malmö: School of Education, No. 30, 1966, (Rev. uppl., 1975).
- Holmberg, I. Effekter av ny läroplan i matematik (Lgr 69) på vissa aspekter av talbegreppets utveckling, kunskaper i matematik och attityder mot ämnet. *Pedagogisk-psykologiska problem*. Malmö: School of Education, No. 236, 1974.
- Horn, R. E. Learner-controlled use of informational retrieval systems. *Programmed Instruction*, IV, 5—12, 1964.
- Hough, J. B. Research vindication for teaching machines. *Phi Delta Kappan*, 1962, 42, 240—242.
- Hough, J. B. & Revsin, B. Programed instruction at the college level. *Phi Delta Kappan*, 1963, 44, 286—291.
- Hughes, J. L. & McNamara, W. J. A comparative study of programed and conventional instruction in industry. *J. appl. Psychol.*, 1961, 45, 225—231.
- Husén, T. (Ed.) *International study of achievement in mathematics*. Vol. I—II. Stockholm: Almqvist & Wiksell, 1967.
- Håstad, M., Svensson, L. & Öreberg, C. *Matematik NU* för lärare på låg- och mellanstadiet. Deltaprojektet, Hermods, 1967 (a).
- Håstad, M., Svensson, L. & Öreberg, C. *Matematik NU* för lärare på högstadiet. Deltaprojektet, Hermods, 1967 (b).
- Härnqvist, K. DBA, *Differentiell begåvningsanalys*. Stockholm: Skand. Testförlaget. 1969.
- Jansson, S. *Matematik för högstadiet*, allmän kurs I och II. Ett undervisningsprogram från Institutet för Programmerad Undervisning, IPU, Göteborg, 1963.
- Jansson, S. Matematikkursens omfång vid individualiserad undervisning. *Skola och samhälle*, 1964, 45, 6—10.
- Johannesson, I. Elevernas inställning till lärare och skolsituationen i olika differentieringsmiljöer. I: Johannesson, I. & Magnusson, D. *Social- och personlighetspsykologiska faktorer i relation till skolans differentiering*. (SOU 1960: 42), Stockholm, 1960.
- Keil, K. A. Erfahrungsbericht über Versuche an höheren Schulen mit einem Programm aus der Raumgeometrie. I: *Praxis und Perspektiven des programmierten Unterrichts*. (Referate des 3. Nürtinger Symposions über Lehrmaschinen.) Quickborn: Verlag Schnelle, 1965, 29—35.
- Kerlinger, F. N. *Foundations of behavioral research. Educational and psychological inquiry*. New York, 1964.
- Lane, G. *Erfarenheter av ett försök med programmerad undervisning i matematik*. Rapporter från Pedagogiska institutionen, Göteborgs universitet, 1963.
- Larsson, I. (Red.) *Individualiserad matematikundervisning: en bok om IMU-projektet*. Malmö: Hermods, 1973 (a).
- Larsson, I. *Individualized mathematics teaching*. Lund: Gleerups, 1973 (b).
- Leith, G. O. M. *A Handbook of Programmed Learning* (2nd ed.). University of Birmingham: Educational Review Occasional Publication, 1966, 1.
- Leith, G. O. M. *Second thoughts on programmed learning*. London: National Council for Educational Technology, 1969.
- Levin, G. R. & Baker, B. L. Item scrambling in a self-instructional program. *J. educ. Psychol.*, 1963, 54, 3.

- Lindquist, E. F. *Statistical analysis in educational research*. New York: Houghton Mifflin Co., 1956.
- Lord, F. M. & Nowick, M. R. *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley, 1968.
- Lovell, K. *The growth of basic mathematical and scientific concepts in children*. London: University of London Press, 1971.
- Lumsdaine, A. A. & Glaser, R. (Eds.) *Teaching machines and programmed learning: A source book*. Washington, D. C.: NEA, 1960.
- Lumsdaine, A. A. (Ed.) *Student response in programmed instruction. A symposium on experimental studies of cue and response factors in group and individual learning from instructional media*. Washington, D. C.: National Academy of Sciences, National Research Council, 1961.
- Lundgren, U. *Pedagogical frames and the teaching process*. Pedagogiska institutionen, Göteborgs universitet, rapport nr 25, 1973.
- Läroplan för grundskolan*. Stockholm: Skolöverstyrelsen, 1962. (Lgr 62.)
- Läroplan för grundskolan*. Stockholm: Utbildningsförlaget, 1969. (Lgr 69.)
- Lövgren, E. *Läsprov A och B*. Stockholm, 1964. (Stencil.)
- Maccoby, N. & Sheffield, F. D. Theory and experimental research on the teaching of complex sequential procedures by alternate demonstration and practice. I: Finch, G. & Cameron, F. (Eds.) *Symposium on Air Force human engineering, personnel, and training research*. Washington, D. C.: National Academy of Sciences, National Research Council, 1958.
- Mager, R. F. On the sequencing of instructional content. I: Davies, I. K. & Hartley, J. (Eds.) *Contributions to an Educational Technology*. London: Butterworths, 1961.
- Mager, R. F. & Clark, P. C. Explorations in student controlled instruction. I: *Psychological Reports*, 1963, 13, 71—76.
- Magne, O. *Matematiksvårigheter*. (Pedagogiska skrifter, 253.) Stockholm: Sveriges Lärarförbund, 1973 (a).
- Magne, O. Den nya matematiken — varför är det så svårt? Och hur kan basfärdigheterna förbättras? *Särtrycksserien*, Malmö: School of Education, No 136, 1973 (b).
- Margolius, G. J. & Sheffield, F. D. Optimum methods of combining practice with filmed demonstration in teaching complex response sequences: Serial learning of a mechanical assembly task. I: Lumsdaine, A. A. (Ed.) *Teaching machines and programmed learning: A source book*. Washington, D. C.: NEA, 1960.
- Markle, S. M. *Good Frames and Bad* (2nd ed.). New York: Wiley, 1969.
- Mechner, F. M. Behavioural analysis and instructional sequencing. I: Lange, C. P. (Ed.) *Programmed Instruction*. Chicago: National Society for the Study of Education, 1967.
- Medley, D. M. & Mitzel, H. E. A technique for measuring classroom behavior. *J. educ. Psychol.*, 1958, 49, 86—92.
- Medley, D. M. & Mitzel, H. E. Measuring classroom behavior by systematic observation. I: Gage, N. L. (Ed.) *Handbook of research on teaching*. Chicago: Rand McNally, 1964, 247—328.
- Mellingsæther, O. A. *IMU. Rapport fra 8. klasse i feltforsøket*. Forsøksrådet for Skoleverket, 1972.
- Mårdberg, B. *Dimensionsanalys för beteendevetare*. Stockholm: Personaladministrativa rådet, 1969.
- New trends in mathematics teaching. Vol. 3*, 1972. UNESCO, Paris, 1973.
- Newkirk, R. L. A comparison of learner control and machine control strategies for computer assisted instruction. I: *Programmed Learning & Educational Technology*. 1973, 10, 82—91.

- Nilsson, L. *Frånvaron i skolan*. Pedagogiska institutionen, Göteborgs universitet, rapport nr 11, 1973.
- Nordisk skolmatematik. *Nordiska utredningar*, 1967: 9. Stockholm, 1967.
- Pask, G. Adaptive machines. I: Bresson, F. & Montmollin, M. (Eds.) *La Recherche en Enseignement Programmé : Tendances Actuelles*. Dunod, Paris, 1969.
- Piaget, J. *The child's conception of number*. London: Routledge and Kegan Paul, 1969.
- Porter, D. Some effects of year long teaching. Machine instruction. I: Galanter, E. (Ed.) *Automatic teaching: The state of the art*. New York, Wiley, 1959, 85—90.
- Pressey, S. L. A simple apparatus which gives test and scores—and teaches. *School and Society*, 1926, 23, 373—376.
- Roe, A. A comparison of branching methods for programmed learning. *J. educ. Res.*, 1962, 55, 407—416.
- Roe, K. V. Scrambled vs. ordered sequence in auto-instructional programs. *J. educ. Psychol.*, 1962, 53, 101—104.
- Romberg, T. A. Current research in mathematics education. *Rev. educ. Res.*, 1969, 39, 473—491.
- Schröter, G. Untersuchung über die Einstellung von Berufsschülern zur programmierten Unterweisung. I: *Praxis und Perspektiven des programmierten Unterrichts*. (Referate des 3. Nürnberger Symposions über Lehrmaschinen.) Quickborn: Verlag Schnelle, 1965, 17—28.
- Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- Silberman, H. F. Characteristics of some recent studies of instructional methods. I: Coulson, B. (Ed.) *Programed learning and computer-based instruction*. New York, Wiley, 1962.
- Skinner, B. F. The science of learning and the art of teaching. *Harvard educ. Rev.*, 1954, 24, 86—97.
- Skinner, B. F. & Holland, J. G. The use of teaching machines in college instruction. I: Lumsdaine, A. A. & Glaser, R. (Eds.) *Teaching machines and programmed learning: A source book*. Washington, 1960, 159—172.
- Starkweather, J. A. Computest: A computer language for individualized testing, instruction, and interviewing. *Psychol. Rep.*, 1965, 12, 227—237.
- Stolurow, L. M. & Walker, C. C. A comparison of overt and covert response in programmed learning. *J. educ. Res.*, 1962, 55, 321—432.
- Stukát, K. G. *Ett försök med programmering av svensk språklära*. Göteborg: IPU, 1963.
- Suppes, P. Plug-in instruction. *Saturday Rev.* 1966, 23, 25—30.
- Tait, K., Hartley, J. R. & Anderson, R. C. Feedback procedures in computer assisted arithmetic instruction. I: *Brit. J. educ. Psychol.*, 1973, 43, 161—171.
- Vanås, E. Räknefärdigheten i klass 15 i några städer. *Aktuellt från Skolöverstyrelsen*, 1953, 6, 311—318.
- Weiss, W., Maccoby, N. & Sheffield, F. D. Combining practice with demonstration in teaching complex sequences: Serial learning of a geometric-construction task. I: Lumsdaine, A. A. (Ed.) *Student response in programmed instruction*. A symposium on experimental studies of cue and response factors in group and individual learning from instructional media. Washington, D. C.: National academy of sciences, National research council, 1961.
- Werdelin, I. *The mathematical ability*. Lund: Gleerups, 1958.
- Werdelin, I. *Sammanställning av resultat av forskning rörande matematikundervisningen*. Linköping: Lärarhögskolan, 1973.
- White, R. T. The validation of a learning hierarchy. I: *Amer. educ. Res. J.*, 1974, 11, 121—136.

- Wilcox, B. The teaching of several tasks using chaining strategies. *Brit. J. educ. Psychol.*, 1974, 44, 175—183.
- Wohlwill, J. F. A study of the development of the number of concept by scalogram analysis. *J. genet. Psychol.*, 1960, 97, 345—377.
- Wohlwill, J. F. & Lowe, R. C. Experimental analysis of the development of the concept of number. *Child Development*, 1962, 33, 153—167.
- Wright, E. M. J. Development of an instrument for studying verbal behaviors in a secondary school mathematics classroom. *J. exp. Educ.*, 1959, 28, 103—121.
- Wright, E. M. J. & Proctor, V. H. *Systematic observation of verbal interaction as a method of comparing mathematics lessons*. St. Louis, Mo.: Washington Univer., 1961.
- Västgötaundersökningen: Andersson, L. E., Norvén, L.-G. & Andreasson, B. Undersökning av kunskaper i numerisk räkning i åk 7. Göteborg: Lärarhögskolan, 1972. (Stencil.)
- Zuckerman, C. B., Marschall, G. R. & Groesberg, S. *Research in the automation of teaching*. Port Washington, New York: U.S. Naval Training Device Center, 1961.
- Öreberg, C. Individualiserad matematikundervisning: Rapport från pågående försök. *Pedagogisk-psykologiska problem*. Malmö: School of Education, No. 8, 1964.
- Öreberg, C. Försök med individualiserad matematikundervisning. I: *Pedagogiska Meddelanden*, 1965, 2 (9), 2—11.

Published by the University of Lund, the Malmö School of Education, and the Society of Psychology and Pedagogy:

STUDIA PSYCHOLOGICA ET PAEDAGOGICA

SERIES ALTERA

1. Edlund, Sven: Diskussionen om begåvningsurvalet under reformations- och stormaktstiden. I. 1947.
2. Agrell, Jan: Nyinrättade professurer inom filosofiska fakulteten i Lund under 1800-talets första hälft. 1949.
3. Smith, Gudmund: Psychological Studies in Twin Differences. 1949.
4. Husén, Torsten: Om innebörden av psykologiska mätningar. Några bidrag till psykometriken metodlära. 1949.
5. Smith, Gudmund: Interpretations of Behaviour Sequences. 1952.
6. Johannesson, Ingvar: Studier av sociala relationer mellan barn i folkskoleklasser. 1954.
7. Kragh, Ulf: The Actual-Genetic Model of Perception-Personality. 1955.
8. Bjerstedt, Åke: Interpretations of Sociometric Choice Status. 1956.
9. Werdelin, Ingvar: The Mathematical Ability. Experimental and Factorial Studies. 1958.
10. Werdelin, Ingvar: Geometrical Ability and the Space Factors in Boys and Girls. 1961.
11. Borg, Gunnar A. V.: Physical Performance and Perceived Exertion. 1962.
12. Lindell, Ebbe: Den svenska rättskrivningsmetodiken. 1964.
13. Bokander, Ingvar: Stereoscopically Induced Perceptual Conflict between Facial Photographs. A Discussion and Summary of Results. 1968.
14. Bjerstedt, Åke: Twelve Years of Educational and Psychological Research in Sweden: A Bibliography. 1968.
15. Larsson, Bernt: Bayes Strategies and Human Information Seeking. 1968.
16. Bjerstedt, Åke: Educational Technology. Instructional Programming and Didakometry. 1972.
17. Löfgren, Horst: The Measurement of Language Proficiency. 1972.
18. Bierschenk, Bernhard: Självkonfrontation via intern television i lärarutbildningen. 1972.
19. Gestrelus, Kurt: Job Analysis and Determination of Training Needs. Examples of Methods Applied to Teacher Trainers. 1972.
20. Sjöbäck, Hans: The Psychoanalytic Theory of Defensive Processes. 1973.
21. Larsson, Inger: Individualized Mathematics Teaching. 1973.
22. Fredriksson, Lennart: The Use of Self-Observation and Questionnaires in Job Analysis for the Planning of Training. 1974.
23. Idman, Pekka: Equality and Democracy: Studies of Teacher Training. 1974.
24. Bjerstedt, Åke: Social Development and Training in School. 1974.
25. Alwall, Ellen: The Religious Trend in Secular English—Irish Schoolbooks 1850—1861 and 1871—1880. 1974.
26. Holmberg, Ingrid: Effects of some trials to improve mathematics teaching. 1975.